32

Artificial Worlds and Agent-Based Simulation

Till Grüne-Yanoff

INTRODUCTION

Agent-based simulations (ABS) are often hailed as a new way of doing social science. By generating 'artificial societies' on the computer, social scientists may generate a new object of study, or a new tool for scientific study. This article investigates these claims by examining in detail the methods of ABS and its uses in the social sciences. It starts out by illustrating the different uses of ABS, and giving a short historical overview of its emergence. ABS is then contrasted with related practices in the social sciences and other disciplines. A brief philosophical discussion of some of its basic concepts follows.

After this overview, the article addresses the novelty claim head on, and discusses it in relation to existing scientific practices. The general position here will be cautious: the author's view is that ABS does not so much constitute a novel scientific practice, but is closely linked to the existing modelling practices in the social sciences. Thus the view of ABS *as* experiments or *as* theory will be treated critically, while the interactions between these practices will be shown to be fruitful. The view of ABS as models, in contrast, will be presented in a more favourable light, but important differences that concern the way of construction, the validation and link to the target system are pointed out. Finally, potential reasons for the notable difference with which different disciplines have adopted ABS will be discussed.

Based on these results, the uses of ABS in the social sciences are investigated and appraised. This discussion will clarify the differences between the various ABS models. In particular, it is shown that uses determine many characteristics of ABS models, in terms of their level of abstraction, their necessary validation and their need for replication. Again, the position presented here will be cautious: against the many enthusiastic claims made about ABS, this article points out the various deficits that many contemporary simulation studies exhibit, particularly when in the business of explanation or forecasting. Instead, the articles points out uses with somewhat lesser profiles, like deliberation support and heuristic application, as important areas where ABS can be fruitfully employed.

USES OF ARTIFICIAL SOCIETIES

To generate a pattern or regularity with the tools of agent-based simulation is to 'situate an initial population of autonomous heterogeneous agents in a relevant special environment; allow them to interact according to simple local rules, and thereby generate – or "grow" – the macroscopic regularity from the bottom up' (Epstein, 1999: 41). The 'agents' referred to are structural entities of the simulation model. They are 'heterogeneous', because the model can specify different attributes for different agents. The simulation updates these attributes solely on the basis of the agents' interaction with each other and their environment. Agents are 'autonomous', because their interactions are determined by individual behavioural rules. 'Macroscopic regularities' concern the regular behaviour of many or all agents in certain environments. Because the simulation only computes the updating of individual agents' attributes, and the macroscopic events are only summaries of the individual changes, the simulation is said to 'grow' the regularities 'from the bottom up'. The tool of agent-based simulation has found many applications in the sciences. For illustrative purposes, a few will be described here.

Uses in the Social Sciences

ABS are used in the social sciences to generate regularities akin to social phenomena. For this reason, ABS employed in the social sciences are often called *artificial societies*. Epstein and Axtell (1996) present a general recipe for their construction in the book *Growing Artificial Societies: Social Science from the Bottom Up*. Their simulation programme *Sugarscape* is a kind of simulation laboratory. Agents move over a two-dimensional rectangular grid. Their existence, their movement and their interaction is conditional on certain agent attributes and regulated by certain rules. By manipulating the initial number and position of agents, and their attributes and rules, different patterns arise on parts or the whole of the grid. Epstein and Axtell interpret these patterns as similar to certain social phenomena. They claim that their artificial society experiences migration, population developments, cultural identities, markets, wealth inequalities, wars and epidemics. Crucially, they claim that they can *generate* such (representations of) social phenomena by the manipulation of agent attributes and rules alone. The social thus *emerges* from the interaction of individuals in their agent-based simulations.

Many social scientists have adopted ABS as useful tools or as interesting objects of study. Today, one finds agent-based simulations of crowd dynamics (Pan et al., 2007), epidemics (Eubank et al., 2004), consumer markets (Moss, 2002), electricity markets (Bunn and Oliveira, 2003), stock markets (LeBaron, 2001; Samanidou et al., 2007), geopolitical change (Cederman, 2002) and even ancient societies (Dean et al., 2000; Axtell et al., 2002). As will be discussed later, the application of ABS in these fields follows a variety of objectives. Sometimes the goal is to be explained what is simulated, as for example, in the case of the ancient society simulation. In some other cases, the goal is to predict outcomes, for example, of strategic withholding behaviour in the electricity market. Finally, in many cases simulations are used to assess the outcomes of different policy interventions, as for example through crowed control or vaccination programmes.

Uses in Other Sciences

ABS have been widely used in disciplines outside the social sciences. First and foremost is ecology, which has produced as many agent-based models as all other disciplines together over the last 30 years (Grimm et al., 2005). Ecological applications range from modelling the spatiotemporal dynamics of beech forests of central Europe (Wissel, 1992; Rademacher, 2005), spatial distribution of animals (Catling and Coops, 2002) to agricultural change (Happe et al., 2006). Another important application is found in behavioural ecology, which studies the behaviour of animal groups, swarm and flocks. The most famous of these is Reynolds 'boid' model (Reynolds, 1987)

Other Artificial Systems

Computer simulations do not only aim at constructing artificial societies (AS), but are also, and perhaps better, known to be used for the construction of artificial intelligence (AI) and artificial life (ALife). It is clear that all these attempts have the same roots (see section on other simulation approaches), however, the branches have developed important differences.

AI is the attempt to build or construct computers to do the sort of things minds can do. Major textbooks define the field as 'the study and design of intelligent agents' where an intelligent agent is a system that perceives its environment and takes actions that maximise its chances of success (Poole et al., 1998: 1; cf. also Nilsson, 1998; Russell and Norvig 2003). Classical AI differs from AS in its focus on top-down processing. In top-down processing, a high-level description of the task (e.g. a goal or a grammar) is employed to start, supervise or guide detailed actions. In bottom-up processing, in contrast, it is the detailed input of the system that determines what will happen next. In classical AI, the top-down approach dominates, while AS exclusively operates on bottom-up processing.

Although connectionist AI works bottom up, it still differs from AS because it cannot alter its fundamental organisation. The behaviour of a connectionist model depends on the local interactions of the individual units, none of which have an overall view of the task. In that respect, it resembles AS. But all a connectionist system can do is to change the weights of different connections; while the fundamental organisation of the connections remains unchanged (Boden, 1996: 4). AS, in contrast, can completely reorganise its own social structure – in fact, it can even cease to exist as a consequence of its own actions.

ALife employs computer simulation to study systems of life, its processes and its evolution. Although the concept of life has no universally agreed definition, certain properties are often associated with it, including autonomy, reproduction and metabolism (Boden, 1996: 1). These concepts give a first distinction of ALife from AS: while ALife studies systems that are autonomous, are reproducing and have a metabolism, none of that can be said of systems studied by AS. Although agents in AS may be autonomous, and may reproduce, artificial societies are not assumed to do so – and of course, they do not metabolise either. Further, ALife is beset with the question of *strong ALife*: whether the entities created in the computer could be genuinely said to be Alife (Ray, 1996; Sober, 1996). A similar question does not exist for AS, which generally is left burdened with less metaphysical baggage.

Other Simulation Approaches

Many different computer simulation techniques are used in the social sciences. ABS must be distinguished from equation-based and micro-simulations. Sometimes, ABS are also distinguished from Cellular Automata (CA), although the difference is not so clear-cut.

Equation-based simulations describe the dynamics of a target system with the help of equations that capture the deterministic features of the whole system. Typical examples of such equation-based simulations are system dynamics simulations, which use a set of difference or differential equation that derive the future state of the target system from its present state. System dynamics simulations are restricted to the macro level: they model the target system as an undifferentiated whole. The target system's properties are then described with a set of attributes in the form of 'level' and 'rate' variables representing the state of the whole system and its dynamics. ABS, in contrast, lack an overall description of the system's macro-properties. Instead of simulating the system's dynamics by numerically calculating the equations that describe the system's dynamics, ABS *generate* the system's dynamics by calculating the dynamics of the system's constituent parts, and aggregating these dynamics into the system dynamics.

Micro-simulations predict the effect of aggregate changes (e.g. taxation changes) on aggregate results (e.g. tax revenue) by calculating the effect of the aggregate change on sub-groups or individuals and then aggregating the individual results. No interaction between groups or individuals is taken into account here; rather, the effect on each group is determined by equations pertaining to this group. Thus, despite its focus on the micro-level, and the subsequent constitution of the macro-result as an aggregate of the micro-level, micro-simulations belong in the equation-based category. What sets ABS apart from such micro-level equation-based simulations is that they model interactions between autonomous agents, thus including a level of *complexity* not existent in equation-based models.

Cellular automata (CA) share this complexity with ABS. CAs consist of cells in a regular grid with one to three dimensions. Every cell has a number of states, which change in discreet time. The states of a cell at a given time period are determined by the states of that same cell and of neighbouring cells at earlier times. The specific kind of these influences is laid down in behavioural rules, which are identical for all cells. A famous example of a CA is Conway's Game of Life (Berlekamp et al., 1982), in which each cell is either 'dead' or 'alive'. 'Dead' cells with exactly three neighbours become 'alive', and 'alive' cells with fewer than two or more than three neighbours die. Conway's Game of Life has attracted much interest because of the surprising ways in which patterns can evolve. It illustrates the way that complex patterns can emerge from the implementation of very simple rules.

When the internal processing abilities of automata are sufficiently high, one speaks about 'agents', and agent-based simulations, not cellular automata. Agents share with CA their autonomy from others' direct control, their ability to interact with others, react to environmental changes and actively shape the environment for themselves and others. In contrast to CA, agents are not fixed on a grid, can change their neighbours, may have multiple relations with different agents and can change on multiple levels. Because the agents' processing abilities are a matter of degree, it is difficult to draw a clear line between CA and ABS.

History

While the history of CA begins in the late 1940s with von Neumann's construction of a self-replicating machine on a paper grid, the beginnings of agent-based models lie with Thomas Schelling's model of segregation (Schelling, 1969, 1971).[1] Schelling originally used coins and graph paper rather than computers, but his models embody the basic concept of agent-based models as autonomous agents interacting in a shared environment with an observed aggregate, emergent outcome. Shortly afterwards, and apparently completely independent of Schelling, the British biologist Maynard Smith and the American chemist George R. Price devised a simulation technique for their novel solution concept of an 'evolutionary stable strategy' (Maynard Smith and Price, 1973; cf. also Sigmund, 2005). This method was very much at odds with the analytic methods used by game theorists at the time, and now stands as one of the earliest examples of an agent-based model. ABS thus had close links with evolutionary game theory and the study of adaptive systems early on and has retained them since (see Chapter 20).

ABS gained serious attraction with the advent of more powerful computers, and the development of agent-based simulation packages like SWARM and Sugarscape in the 1990s. It is noteworthy that the majority of applications in the social sciences can be found in sociology (cf. Chapter 18), while economics has been rather slow in adopting these computational methods (for an overview of the development in economics, see Tesfatsion, 2006).

CONCEPTUAL FOUNDATIONS

In this brief overview of ABS and its related fields, a number of concepts were used to characterise and differentiate ABS. However, these notions are often philosophically problematic. In the following, some of these problems are addressed.

Agents and Agency

The central difference between CA and ABS is the presence of autonomous agents in the latter (see also Chapters 8 and 14). However, cells in a CA are also often said to enjoy a degree of autonomy, so that the central difference must lie in the agency of ABS elements. But what does that mean?

Agency and action are related to the intuitive distinction between the things that merely happen to people — the events they suffer — and the various things they genuinely do. These latter doings are the actions of the agent. Philosophers have discussed at length about the nature of action, in particular, what distinguishes an action from a mere happening or occurrence. Yet most of these discussions are too fine-tuned for the current purpose. Here, it is sufficient to point out that an often used criterion for agency is goal-directedness: In artificial intelligence, for example, an *intelligent agent* is a unit which observes and acts upon an environment and directs its actions towards achieving certain goals (e.g. Russell and Norvig, 2003). Such agents may be very simple: a thermostat is an intelligent agent, for example, which measures the environment and interferes in the heating mechanisms, with the goal of maintaining a constant temperature in the room. Further, the purposeful behaviour of animals constitutes a low-level type of purposive action. When a spider walks across the table, the spider directly controls the movements of his legs, and they are *directed at* taking him from one location to another. Those very movements have an aim or purpose for the spider, and hence they are subject to a kind of teleological explanation (Frankfurt, 1978).

Most ABS, in contrast, fail to endow their agents with a purpose. Agents in ABS commonly behave according to homogenous behavioural rules. They may differ in their attributes (e.g. age, sex, home address, occupation) and may exhibit different behaviour to the same situation due to these differences in attributes, yet these attributes are not distinguished into purposive and non-purposive ones. In the Anasazi simulation (Dean et al., 2000), for example, there is no difference when an agent moves place of settlement to when an agent dies of old age. Both are direct functions of environmental variables, and the settlement decision is – contrary to one's intuitive identification as a purposeful action – not mediated by any kind of utility function of the agent. Similarly in the smallpox simulation (Eubank et al., 2004): there is no difference between the agents' going to work and them getting infected. Although many would say that the first is a purposeful action, and the second is not, the ABS does not make a difference between them. This lacking attribution of any goal or utility function to the ABS agents makes it difficult to subject them to any teleological explanation; and hence the agency of ABS agents is somewhat insecure.

Artificiality

The results of ABS in the social sciences are often called 'artificial societies'. The nature of these societies' artificiality is ambiguous. Depending on its interpretation, not all ABS in the social sciences may be artificial societies, or artificial societies may be the results not only of ABS. At least three different interpretations can be distinguished: artificiality as (1) non-realisticness, (2) constructedness or (3) computer-generated.

Artificial societies may refer to the classes of unrealistic simulations. In some sense, all ABS-generated societies are artificial in this sense: they are codes on a computer platform and not societies in the world. Yet some of these ABS yield societies realistic at least in some respects, in that they closely resemble *aspects* of human societies. The Anasazi ABS (Dean et al., 2000), for example, is claimed to resemble the housekeeping, reproductive and migration behaviour of actual ancient people on the micro-level, and it is also claimed to resemble its population and settlement dynamics on the macro-level. In contrast to this, other ABS are not intended as simulations of human societies at all. Doran (1998), for example, simulates a society in which agents have perfect foresight (i.e. they know what will befall them in the future). Such a society may be called artificial in the sense that it is largely based on assumptions known to be false in the real world. While such a distinction would have certain merits, it would exclude many ABS, despite the fact that these ABS are commonly called artificial societies.

It may therefore be more adequate to interpret 'artificial' as 'constructed' or 'man-made'. Artificial societies are the result of modelling and programming efforts, while real societies have evolved without the intentional input of a designer or constructor. Such a distinction, however, would include under the term artificial societies social phenomena created in the laboratory, as for example produced by experimental economists. This does not seem to comply with standard usage of the term.

One may therefore be forced to narrow the interpretation further to computer simulations. Artificial society refers to processes *in silico*, which model social phenomena, as opposed to material exemplars of these processes themselves. However, the material basis of social phenomena could be challenged. While this issue has not arisen in the context of artificial societies so far, a lively debate exists with respect to artificial life. The strong Alife position claims that life is a process that can be abstracted away from any particular medium. Defenders of such a position avoid the term artificial and instead prefer 'synthetic biology' (Ray, 1996). In a similar vein, it could be argued that sociality is a process independent of its material basis, and that computer-generated interactions between agents constitute social phenomena in the same way as interactions of real-world agents.

Emergence

ABS generate macroscopic patterns from 'the bottom up' – from the interactions of many microscopic agents. It is often said that these macroscopic patterns or properties *emerge* from the microscopic ones. The term 'emergence' has a long history, which unfortunately makes its use more difficult. Some conceptual clarification is necessary.

In early twentieth century philosophy, emergence came to denote a metaphysical position on higher-level properties. Emergent properties, it was claimed, 'arise' from some lower-level properties and yet are 'irreducible' to them. The positions varied in the way properties arose and were said to be irreducible. Although different in detail, both Mill (1843) and Broad (1925) claimed that these properties are associated with irreducible higher-level causal powers; while Alexander (1920) claimed that although the emergent properties were primitive, their causal powers do not supersede more fundamental interactions. It is controversial whether Alexander's proposal is coherent as an ontological position, while Mill's and Broad's imply *downward causation*, which sits uncomfortable with many. According to downward causation, the emergent property's causal influence is irreducible to the micro-properties from which it emerges; it bears its influence in a *direct* downward fashion, *not* through the aggregation of its micro-level powers (for a defence, see O'Connor, 1994).

One way to avoid downward causation is to re-interpret emergence as an epistemic, not an ontological notion. Emergent properties then are characterised in terms of limits on human knowledge of complex systems. A property is thought of as emergent, accordingly, if it could not have been predicted from knowledge of features of its parts. More specifically, it could not have been predicted without the help of simulation (Bedau, 1997: 378). Such a concept of *weak emergence* is conceptually unproblematic, but does not capture the metaphysical spirit of previous notions.

ABS IN THE CONTEXT OF SCIENTIFIC PRACTICE

Simulation, Models and Theory

Simulations rest on models (see Chapter 28). Without models of the networks through which epidemics spread, or the landscape on which the Anasazi settle, the examples of the previous section could not have simulated anything. But while it is uncontroversial that models are important and possibly constitutive elements of simulations, it is less clear whether simulations themselves can be treated as models. Of course, in a colloquial sense, models and simulations are not properly delineated. For example, most economists think of Schelling's (1971) checkerboard model as a model, while it is also considered to be one of the earliest examples of an agent-based simulation. However, a number of differences can be identified.

One difference is the temporal dimension of simulations. Scientists often speak about a model 'underlying' the simulation. The recent smallpox infection simulation of Eubank et al. (2004), for example, is based on a model of Portland, OR, consisting of around 181,000 locations, each associated with a specific activity, like work, shopping, school and maximal occupancies. Additionally, each model inhabitant is characterised by a list of the entrance and exit times into and from a location for all locations that person visited during the day. This huge database was developed by the traffic simulation tool TRANSIMS, which in turn is based on US census data. When speaking about the model underlying the simulation, people often have such a static model in mind. The simulation itself proceeds by introducing a hypothetical 'shock' into the system (in this case, a number of infected inhabitants) and then observing how the infection spreads through the population. This dynamic aspect is often not included when people speak about the underlying model. This may be a sensible distinction, as the dynamic aspect of the simulation makes various diachronic stability assumptions that were not included in the static model (Grüne-Yanoff, 2010). Of course, the dynamic aspects may be referred to as the 'dynamic model' which includes the 'static model', yet common practice in such cases seems to be that the 'static model' is referred to as the 'underlying model'.

A second difference lies in the methods by which models can be analysed. The common way that mathematical models in the natural science or in economics are analysed is to find a solution to the set of equations that make up the model. For this purpose, calculus, trigonometry and other mathematical techniques are employed. Being able to write down the solution this way makes one absolutely sure how the model will behave under any circumstances. This is called the *analytic solution*, because of the use of analysis. It is also referred to as a closed-form solution.

However, analytic solutions work only for simple models. For more complex models, the maths becomes much too complicated. Instead, the model can be 'solved' by simulation. For, say, a differential equation that describes behaviour over time, the numerical method starts with the initial values of the variables, and then uses the equations to figure out the changes in these variables over a very brief time period. A computer must be used to perform the thousands of repetitive calculations involved. The result is a long list of numbers, not an equation. Appropriately presented, numerical simulation is often considered a 'solution' of the model.

Some proponents of simulation have argued that for every computation, there is a corresponding logical deduction (Epstein, 1999), hence from a technical standpoint, deductive modelling is but a special case of simulative modelling. However, this claim neglects important epistemic and psychological differences. As Lehtinen and Kuorikoski (2007) point out, economists largely shun simulations for epistemic and understanding-related reasons. They explain this observation by arguing that economists place a high value on the *derivation* of an analytical result, based on their belief that the cognitive process of solving a model constitutes the understanding of the model. In most simulations, the computer is a necessary tool: humans could not, even if they wanted to, perform the computations needed. The derivation of results in these simulations is outside of the reach of human agents. They leave the solution process, in the words of Paul Humphreys, 'epistemically opaque'. This opaqueness makes economists shun simulation when they seek understanding from the analytic solution process itself. It also constitutes an important difference between standard (analytically solvable) models and simulations.

To summarise, simulations differ from models mainly in their temporal expansion (and sometimes also in their representation of a temporal process) as well as in their epistemic opacity.

Simulations versus Experiments

Another perspective on simulations links them to experiments (see Chapter 30; see also Dowling, 1999; Rohrlich, 1991: 507). Because simulations are typically based on calculations that are intractable, the results of a simulation cannot be predicted at the time when the simulation is constructed or manipulated. This allows seeing the simulation as an unpredictable and opaque entity, with which one can interact in an experimental manner. However, the legitimacy of a computer simulation still relies on the analytic understanding of at least the underlying mathematical equations, if not the computation process itself. Thus the experimental approach to simulations consists in a strategic move to 'black-box' (Dowling, 1999: 265) the known programme, and to interact 'experimentally' with the surface of the simulation.

Whether this observation suffices to subsume simulations under experiments, remains an open question. Most scientists agree that simulations have experimental moment, but hasten to add a qualifier, for example, that simulations are 'computer experiments'. Along these lines, many philosophers of science have pointed out that despite their experimental moment simulations differ from experiments in important ways.

The first argument for such a difference points to a perceived difference in the similarity relations of experiments and simulations to their targets. Gilbert and Troitzsch (1999: 13), for example, argue that in a real experiment, one controls for the actual object of interest – while in a simulation one is experimenting with a model rather than the object itself. Following a thought of Herbert Simon (1969), Guala (2005: 215) addresses a similar issue, arguing that in a real experiment, the same material causes are at work as those in the target system; while in simulations, not the same material causes are at work, and the correspondence between the simulation and its target is only abstract and formal.

Parker (2009) contradicts these claims. She points out that the use of simulations in what she calls 'computer simulation studies' involves intervention, just as laboratory experiments do. Computer simulation studies intervene on a material system, namely the programmed computer. Such studies are thus material experiments in a straightforward sense.

The second argument for the difference between experiments and simulations points out the different epistemological challenges that experiments and models face. Morgan (2003: 231) argues that they differ in their 'degree of materiality', and that this makes experiments epistemically privileged compared to simulations. One can argue for the external validity of laboratory experiments by pointing out that they share 'the same stuff' with their targets. Simulations, however, only have a formal relation to their targets, which makes establishing their external validity that much harder. Note that this argument draws on the ontological difference identified above; yet Morgan stresses the epistemological implications of these differences, and does not claim that simulations are otherwise fundamentally different from experiments.

Winsberg (2009) offers another version of this epistemological argument. Instead of drawing on the make-up of simulations, he argues that the justification for the claim that a simulation stands for a target rests on something completely different from a similar justification for experiments. The justification for simulation rests on our trust in the background knowledge that supports the construction of the simulation, in particular principles deemed reliable for model construction. The justification for experiments, in contrast, relies on a variety of factors, the most prominent maybe being that experimental object and target are of the same kind. Thus, Winsberg denies, *pace* Morgan, that experiments are epistemically privileged, but insists that the knowledge needed for a good simulation is different from the knowledge needed for a good experiment.

The Novelty Claim

Related to the above discussions is the question whether and to what extent simulation poses a novelty for philosophy of science. While it is obvious that simulation has brought many innovations to science, it is more controversial whether simulation poses new problems for the philosophy of science. Schweber and Wächter (2000) for example, suggest that the widespread use of simulation in the sciences constitutes a 'Hacking-type revolution'. By this they mean that modelling and simulation have achieved a qualitatively new level of effectiveness, ubiquity and authority. Consequently, new problems arise for philosophy of science. Rohrlich (1991: 507) argues that computer simulations require a new and different methodology for the physical sciences. Humphreys (1991: 497) agrees that computer simulations require 'a new conception of the relation between theoretical models and their application'. He advances similar arguments in his 2004 book. Finally, Winsberg (2001: 447) claims that 'computer simulations have a distinct epistemology'.

Against these novelty claims, others have argued that simulations are very similar to traditional tools of science, and do not constitute a revolution in the principles of methodology (Stöckler, 2000). To understand these arguments better, it is helpful to analyse in which way simulations are supposed to pose new problems for the philosophy of science. Frigg and Reiss (2009: 595) present the following list of purportedly novel problems:

1. *Metaphysical:* Simulations create some kind of parallel world in which experiments can be conducted under more favourable conditions than in the 'real world'.
2. *Epistemic:* Simulations demand a new epistemology.
3. *Semantic:* Simulations demand a new analysis of how models/theories relate to concrete phenomena.
4. *Methodological:* Simulating is a sui generis activity that lies 'in between' theorising and experimentation.

Against (1) Frigg and Reiss argue that the parallel world claim already has been made with respect to standard deductive models (cf. Sugden, 2000). Against (2) they argue that the issues with simulation are part of the larger problem from where (complex) models get their epistemic credentials. Against (3) they argue that first, simulations do not clash with either the semantic or the syntactic view, and second, that the dynamic aspect of simulation is not new. Against (4) they argue, first, that simulation does not have an 'in-between status' with respect to its reliability, but that, second, other interpretations of simulations being 'in-between' – like being a hybrid or a mediator – are not new and have been claimed for models already.

Against this sceptical position, Humphreys (2009) argues for the truth of at least (2) and (3). He argues that the epistemic opacity of simulations and their dynamic aspects are new features that are not sufficiently captured by existing accounts of philosophy of science. In addition, he claims that the application process of the simulation to the real world requires a new conceptual framework, and that the limitations of what is computable and hence simulatable in a given time have important implications for the philosophical debate as well.

In this debate, a lot obviously depends on how simulation is defined (cf. section on Uses in the Social Sciences). Frigg and Reiss prefer a more abstract account of simulation that is not strongly differentiated from models, while Humphreys prefers an account that is clearly embedded in the programming and computer implementation of simulation. We feel that both positions have their merits. The sceptical position helps one not get too distracted when trying to explain how modern science works: it avoids the abandonment of central but enduring problems for novel but possibly superficial problems of current practice. The novelty position takes the actual practices of scientists very seriously, as have previous philosophers of science (e.g. Kuhn or Hacking). We believe that the debate between these two factions will not be resolved soon. Many of the problems of more traditional practices of science, which the sceptics claim can account for simulation as well, have not been given a satisfactory solution so far. Whether there are special problems of simulation remaining may only come into high relief once these more general issues have been adequately addressed, and the relevance of their answers for simulations explored.

THE SCIENTIFIC USES OF SIMULATIONS

The sciences use simulations for multiple purposes. In this section, we first explicate how scientists pursue their aims with the help of simulations, and secondly point out the conditions necessary to justifiably pursue these uses with simulations.

PREDICTION

A prediction is a claim that a particular event will occur (with certain probability) in the future (see also Chapter 34). A simulation may predict a phenomenon without explaining it. For example, a model bridge may show that a design will work without explaining why it will work. A model car's performance in a wind tunnel simulation may indicate the car's wind resistance without explaining its wind resistance. However, such cases might be restricted to material simulations: one may be able to successfully exploit the material causes operating in such a simulation for predictive purposes, without being able to identify these causes, and hence without being able to explain why the system operates in the way it does. In non-material simulations, in particular in computer simulations, one has to explicitly construct the principles governing the simulation. Claiming that such a simulation could predict without explaining would then raise the 'no miracles' argument: predictive success would be miraculous if the simulation and its underlying principles did not identify the actual causes at work in the real system. Full *structural validity* of the model – that is, the model not only reproduces the observed system behaviour, but truly reflects the way in which the real system operates to produce this behaviour – vouches for both predictive and explanatory success.

Yet there are different ways in which simulations are based on 'underlying principles'. The simplest is the case in which the simulation is based on natural laws. Take for example vehicle crash simulations. A typical 'first principle' crash simulation takes as input the structural geometry of a vehicle and the material properties of its components. The vehicle body structure is analysed using spatial discretisation: the continuous movement of the body in real time is broken up into smaller changes in position over small, discrete time steps. The *equations of motion* hold at all times during a crash simulation. The simulation solves the system of equations for acceleration, velocities and the displacements of nodes at each discrete point in time, and thus generates the deformation of the vehicle body (cf. Haug, 1981).

Such 'first principle' simulations were built to predict effects of changes in vehicle composition on the vehicle's crash safety. They analyse a vehicle 'system' into its components and calculate the behaviour of these components according to kinematic laws (partly expressed in the equations of motion). Because the computational generation of the behaviour strictly adheres to the causal laws that govern the behaviour in reality, the generation also causally *explains* it.

The builders of crash simulations are in the lucky position that the generated events match the findings of empirical crash tests very precisely, while their models are fully based on laws of nature. This is often not the case. One reason may be the absence of true generalisable statements about the domain of interest. Take for example Coops and Catling's (2002) ecological simulation. Their aim is to predict the spatial distribution and relative abundance of mammal species across an area in New South Wales, Australia. They proceed in the following steps. First, they construct a detailed map of the area indicating for each pixel the 'habitat complexity score' (HCS), which measures the structural complexity and biomass of forested vegetation. This map is estimated from the relationship between HCS observed from selected plots and aerial photographs taken of the whole area. Second, they estimate a frequency distribution of HCS for each selected plot. From this they predict the HCS of each pixel at any time period. This prediction in effect constitutes a simulation of the HCS dynamics for the whole area. Finally, they estimate a linear regression model that links HCS to spatial distribution and relative abundance of the relevant mammal species. Based on this model, they simulate the dynamics of the mammal population throughout the area.

Clearly, Coops and Catling cannot base their simulation on natural laws, because there aren't any for the domain of phenomena there are interested in. Instead, their research paper has to fulfil the double task of estimating general principles from empirical data, and then running the simulation on these principles. Understanding this also makes clear that the main predictive work lies in the statistical operations, i.e. the estimations of the HCS frequency distributions and the linear regression model. The simulation of the HCS dynamics is a *result* of the HCS frequency estimations. It then helps provide the data for the linear regression model; but it can only do so (and one would accept the data it provides as evidence only) if the HCS frequency distributions were estimated correctly. The predictive power of the simulation thus clearly depends on the principles used in it, and the validity of these principles seems not very secure in this case.

Another reason for failing to incorporate independently validated principles is that many simulations do not successfully match the target events or history when relying solely on laws, even if those laws are available. Take for example the following case from climate research (described in Küppers and Lenhard, 2005). In 1955, Norman Phillips succeeded in reproducing the patterns of wind and pressure of the entire atmosphere in a computer model. Phillips used only six basic equations in his model. They express well-accepted laws of hydrodynamics, which are generally conceived of as the physical basis of climatology.

Phillips' model was a great success, because it imitated the actually observed meteorological flow patterns well. But the model also exhibited an important failure: the dynamics of the atmosphere were stable only for a few weeks. After about four weeks, the internal energy blew up, and the system 'exploded' — the stable flow patterns dissolved into chaos.

Subsequent research searched for adequate smoothing procedures to cancel out the errors before they could blow up. This strategy was oriented at the ideal of modelling the true process by deriving the model from the relevant laws in the correct fashion. Instabilities were seen as resulting from errors – inaccurate deviations of the discrete model from the true solution of the continuous system.

The decisive breakthrough, however, was achieved through the very different approach of Akio Arakawa. It involved giving up on modelling the true process, and instead focused on imitating the dynamics. To guarantee the stability of the simulation procedure, Arakawa introduced further assumptions, partly contradicting experience and physical theory. For example, he assumed that the kinetic energy in the atmosphere would be preserved. This is definitely not the case in reality, where part of this energy is transformed into heat by friction. Moreover, dissipation is presumably an important factor for the stability of the real atmosphere. So, in assuming the preservation of kinetic energy, Arakawa 'artificially' limited the blow-up of instabilities. This assumption was not derived from the theoretical basis, and was justified only by the results of simulation runs that matched the actually observed meteorological flow patterns over a much longer period than Phillips' model.

This last example requires us to be more precise when talking about the validity of a model. Structural validity we encountered before: it requires that the model both reproduces the observed system behaviour and also truly reflects the way in which the real system operates to produce this behaviour. But Phillips' model obviously violates structural validity, and still seems to be successful at predicting global weather. In that case, we must speak of *predictive validity*, in which the simulation matches data that was not used in its construction. (One may add that Coops and Catling's 2002 simulation may not be predictively but *replicatively valid*: it matches data already acquired from the real system). By distinguishing structural and replicative validity, we admit that some simulations may predict but do not explain.

Explanation

Agent-based simulations are often claimed to be explanatory (Axtell et al., 2002; Cedermann, 2005; Dean et al., 2000; Epstein, 1999; Sawyer, 2004; Tesfatsion, 2006). Often these claims are ambiguous about how agent-based simulations are explanatory, and what they explain. In the following, we discuss three possible accounts of what kind of explanations ABS may provide (see also Chapters 13 and 33).

Full Explanations

Some simulations are claimed to explain concrete phenomena. Such singular explanations purport to explain why a certain fact occurred at a certain time in a certain way, either by providing its causal history, or by identifying the causal relations that produced it. For example, Dean et al. (2000) claim that by simulating actual population dynamics and settlement densities of the Anasazi, they manage to explain these population dynamics.

Although the simulation matches data during most of the period studied, it does not match data at the period's end. Dean et al.

conclude that some factor outside the simulation influenced population and its distribution at that time. They conjecture that some households left the valley because of social ties to other households leaving the valley and not because potential maize production was not enough to sustain them.

Thus, by the author's own account, the simulation fails as a full explanation of the particular Anasazi history. It omits, besides social pull, social institutions and property rights. It may nonetheless yield a partial explanation that treats some explanatory factors, such as maize production, and controls for other explanatory factors, such as social pull. It may control for an explanatory factor by, say, treating a period during which that factor does not operate. Elaboration of the simulation may add explanatory factors, such as social pull, to extend the simulation's range and make its explanation more thorough. The next section further explores simulations' power as partial explanations of particular phenomena.

However, as Grüne-Yanoff (2009 argues, it is unlikely that this history could ever be explained via simulation, as it is unlikely that the underlying model could ever be sufficiently validated. Instead of providing full or partial explanations of particulars, simulation may only provide *possible explanations*. Such possible explanations, which will be discussed in 4.3.3, may help in the construction of actual explanations, but do not constitute actual explanations themselves.

Partial Explanation

A partial explanation describes the operation of some factors behind a phenomenon's occurrence. This requires the model to successfully *isolate* these explanatory factors (Mäki, 1994). For a partial explanation, each assumption must control for an explanatory factor, or else the theorem's results must be *robust* with respect to variation in the assumption. Yet it turns out that many ABS fail to be robust.

Lacking robustness is a widespread problem for the success of partial explanations with simulation studies. Take for example Huberman and Glance (1993), who examine simulations of generations of players in Prisoner's Dilemmas. The simulations use cellular automata, with cells located in a square. One simulation treats time as discreet and has all cells update at the same time to produce the next generation. Another simulation, more realistically, treats time as continuous so that at a moment at most one cell updates to produce an offspring. Suppose that both the synchronous and the asynchronous simulations begin with the same initial conditions: a single defector surrounded by cooperators. The synchronous simulation maintains widespread cooperation even after 200 rounds, whereas the asynchronous simulation has no cooperation after about 100 rounds. Cooperation is not robust with respect to the updating's timing in these simulations. So unless timing is an explanatory factor in the world and not just an artifact of the simulation, a simulation that generates cooperation using synchronous updating does not yield a partial explanation of cooperation.

Consequently, proponents of the explanatory value of a simulation show that the simulation robustly generates the target phenomenon's representation. That is, the simulation generates the phenomenon's representation over a wide range of variation in the simulation's unrealistic assumptions. The robustness may be with respect to variation in initial conditions, dynamical laws, or values of the simulation's parameters. Justin D'Arms et al. (1998: 89–92), for example, use robustness as a guideline for assessment of simulations and models of adaptive behaviour. They say that a result is robust if it is achieved across a variety of different starting conditions and/or parameters. They take robustness as necessary but not sufficient for a successful simulation.

However, a model's robustness with respect to all assumptions is neither necessary nor sufficient for a phenomenon's partial explanation. A partial explanation requires robustness with respect to variation in assumptions that introduce features irrelevant to the model's target phenomenon. Altering those assumptions should not make a difference to the model's results. In contrast, robustness need not hold with respect to assumptions that control for explanatory factors. In fact, a good model, as it becomes more realistic by incorporating more explanatory factors, does not robustly yield the same results. When it is completely realistic, it exhibits a limited type of robustness. It steadfastly yields its target phenomenon as the model's parameters vary in ways that replicate the phenomenon's natural range of occurrence. Thus, a partial explanation requires only limited robustness, namely, robustness with respect to variation in assumptions that do not control for explanatory factors.

Potential Explanation

We have argued that many simulations often do neither fully nor partially explain any particular phenomenon. Nevertheless, many authors of simulation studies claim that their simulations are in some way explanatory. It may therefore make sense to expand the notion of explanatoriness to include not only full or partial explanations, but also potential explanations. A model or theory may be considered a potential explanation if it shares certain properties with actual explanations, but where the *explanans* is not true (cf. Hempel, 1965: 338). In that sense, simulations may be potential explanations, or as some simulation authors prefer, 'candidate explanations' (Epstein, 1999: 43), and hence may be considered explanatory.

Emrah Aydinonat (2008) offers a good example of such reasoning. He argues that Menger's theory of the origin of money, and more recent simulations building on Menger's work, are partial potential explanations.

Carl Menger (1982) investigated the question how money arose as a medium of exchange. His question was theoretical in that it asked for the general underlying causes for the origin of money, and not for the causal history of any particular instance of money. Envisioning a world of direct exchange, Menger postulated that some goods are more saleable than others, depending on properties like their durability, transportability, etc. Self-interested economic agents, he then argued, would tend to purchase the most saleable good, even if they do not need it, in cases where they cannot directly exchange their goods for goods that they do need. Because everyone would gravitate towards the most saleable good in the marketplace in such situations, it is that good that emerges as the medium of exchange – as the unintended consequences of economising agents.

Aydinonat admits that Menger's model neglects many institutional particularities, and in general is not able to verify its assumptions. It thus cannot offer a full or partial explanation. However, he argues that 'Menger's conjecture alerts us to certain explanatory factors that *may have been* important in the development of a medium of exchange' (Aydinonat, 2008: 48, my emphasis). In particular, Menger's model identifies *some* factors, not all; hence his model offers only a partial explanation. Furthermore, the model identifies only possible factors, not actual ones; hence it offers only a potential explanation.

Many authors have since tried to develop Menger's model further. As an example, take the simulation study by Marimon et al. (1990); they model the trade interactions of three types of agents in the population. Each type consumes a different good, which she does not produce herself. To be able to consume, the agents have to exchange with others. Yet each agent can only store one kind of good, and storage costs for a specific kind of good depends on the type of agent who stores it. In the simulation, agents are matched pairwise at random, offer their goods simultaneously, and decide whether to accept the trade offer. Offers of an agent's consumption goods are always accepted. But if they are not offered their consumption goods, they have to decide whether to accept a good they cannot consume. Agents know a menu of behavioural rules (including 'accept if storage costs are low', 'accept if other agents accept', etc.) and attach strength to each rule. This strength index determines how probable it is that an agent chooses a certain rule. After each round, agents update the strength index according to the success of the rule used.

Marimon et al. find that under specific conditions, the population converges on an equilibrium where every agent prefers a lower-storage-cost commodity to a higher-cost-commodity, unless the latter is their own consumption good. Thus, they show that under these conditions, a medium of exchange emerges as an unintended consequence of the agents' economising behaviour. However, they also find that this convergence is rather sensitive to the initial conditions. Aydinonat therefore concludes that the simulation 'teaches us what we may consider as possible under certain conditions. Yet they do not tell us whether these conditions were present in history or whether there are plausible mechanisms that may bring about this possibility' (Aydinonat, 2008: 112). The simulation offers neither a full nor a partial explanation of the origin of money. But it makes more precise the possible worlds in which Menger's conjecture holds; it specifies in precise detail some environments, and some sets of causal relations under which a medium of exchange emerges. In this sense, the simulation may be considered progress with the possible explanation offered by Menger.

In a similar vein, one may consider the Anasazi simulation progress with possible explanation of the population dynamics. Yet what does the progress consist in? What distinguishes serious contenders for such possible explanations from mere fantastic constructs? Hempel had the formal rigor of the Deductive-Nomological account to fall back onto when referring to the 'other characteristics' of an explanation. But in the age of simulation, indefinite numbers of potential explanations can be produced. With so many possible causes identified, simulation may confuse instead of clarify, and reduce understanding instead of improving it.

One problem, Grüne-Yanoff (2009) argues, may lie in the focus on causes and mechanisms. Aydinonat, for example, claims that simulations 'try to explicate how certain mechanisms … may work together' (2008: 115). Yet these simulations operate with thousands of agents, and indefinitely many possible mechanisms. Identifying a single set of possible mechanisms that produce the explanandum therefore does not, *pace* Aydinonat improve the *chances* of identifying the actual mechanisms. The numbers of possible mechanisms is just too large to significantly improve these chances.

Instead, Grüne-Yanoff suggests that a simulation run offers an instance of the simulated system's *functional capacities* and its functional organisation. Functional analysis shows how lower-level capacities *constitute* higher-level capacities. The capacity of the Anasazi population to disperse in times of draught, for example, is constituted by the capacities of the household agents to optimise under constraints, and their capacity to move. The dispersion *is* nothing but the individual movings. Yet there are many different household capacities that constitute the same higher level capacity. The role of simulation studies, Grüne-Yanoff (2009) argues, is not to enumerate possible household capacities (or mechanisms), but to explore the system's possible functional organisations under which different sets of household capacities constitute higher-level capacities, and hence the 'working' of the whole system. This is in line with current practice. Reports of simulations do not offer comprehensive lists of possible mechanisms that produce the explanandum. Rather, they offer one or a few selected settings, and interpret these as instances of how the system may be functionally organised in order to yield the explanandum. Occasionally, they also conclude from such singular simulation settings that the simulation is not correctly organised and that additional functional components are needed. In the Anasazi case, for example, the authors conclude that additionally push and pull factors are needed. For this reason, it may be preferable to think of simulations as providing potential *functional* instead of potential *causal* explanations.

Policy Formulation

Simulations have long been used to support policy formulation. Drawing on economic theory, Jan Tinbergen constructed a macroeconomic model of the Dutch economy. It led to simulations of six policies for alleviating the Great Depression. Because of the results of the simulations, Tinbergen recommended that the Dutch government abandon the gold standard, which it did.

Today, agent-based models are widely used to simulate the impact of external shocks on complex social phenomena. For example, a number of recent papers have investigated how a smallpox epidemic would spread through a population, and how different vaccination policies would affect this spread. Some of these simulations stay on a relatively abstract level, while others become incredibly detailed and in fact purport to simulate the population behaviour of a whole city (Eubank et al., 2004, who simulate Portland, OR) and even a whole country (Brouwers et al., 2006, who simulate Sweden). Authors of such simulations, in particular from the latter category, often give policy advice based on the simulation results alone.

What kind of policy decisions can be made of course depends on the validity of the simulation (Grüne-Yanoff, 2010). If correct predictions can be made on the basis of the simulation, a straightforward utility maximisation or cost–benefit analysis can be performed. But with most ABS, such point-predictions are out of reach. Instead, ABS at best offer possible scenarios, and allow weeding out certain scenarios as inherently inconsistent or not co-tenable (Cederman, 2005). The goal of simulation studies then is *exploratory modelling*, in which researchers run a number of computational experiments that reveal how the world would behave if the various conjectures about environments and mechanisms were correct.

The results of exploratory modelling are sets or ensembles of possible worlds. This leads to the question how such resulting sets of scenarios can be used as the basis of policy decisions. If the parties to the decision do not know the probabilities of the models in the ensemble, situations of 'deep uncertainty' arise (Lempert, 2002: 7309–7313). Under deep uncertainty, models of uncertain standing produce outcomes with uncertain relevance. Instead of predicting *the* future of the system with one model or with a set of probabilistically weighted models, simulations only yield a 'landscape of plausible futures' (Bankes et al., 2001: 73).

How can the policy maker base her decisions on such a set? Two different strategies have been discussed. The first focuses on worst-case scenarios, against which policies should be hedged. This approach is similar to the maximin decision rule: the policy maker chooses that policy that maximises the minimal (worst) outcome. The second approach pays equal attention to all models, and chooses that policy which performs relatively well, compared with the alternatives, across the range of plausible futures. If 'performs relatively well' is interpreted as performing well against a set minimal threshold, then this approach is similar to the satisficing decision rule: the policy maker sets a threshold in the light of the specific policy goals, and then evaluates the different policy alternatives by their performance in a sufficiently large number of simulation runs.

Both maximin and satisficing are very sensitive to the number of models considered. The wider the scope, the more likely the inclusion of some outlandish terrible future, which will affect maximin choice. Similarly, the wider the scope, the more likely the inclusion of some outlier below the threshold, which will affect satisficing choice. Given the uncertain status of many model specifications, exploratory modelling is prone to such misspecifications. This leads to the question how the scope of the model ensemble can be constrained.

Grüne-Yanoff (2010) argues that neither references to the actual world, nor references to intuitions are sufficient to appropriately restrict the scope of model ensembles. Only through integrating the simulation ensemble under a theory does exploratory modelling gain sufficient systematicity. In such a setting, simulations would unpack the implications of their theoretical hypotheses. If implications are found untenable, the authors can go back to the theory, which provides constraints on how alternative hypotheses can be constructed. Yet current modelling practice rarely follows this approach. For this reason, the usefulness of exploratory modelling for policy formation is not entirely clear.

CONCLUSION

In this article, we argued that agent-based simulation is an important new tool for the social scientist. Although it shares many features with both models and experiments, its dynamic aspects, its ability to compute vast amounts of data, and its epistemic opacity are novel features that set it apart from other scientific tools. This novelty leads to a number of potentially new uses in the sciences. Yet the conceptual foundations for these new employments are still shaky. In particular, we pointed out the potential, but also the difficulties of explaining with simulations, and of supporting policy advice. We hope that this article helps sharpen the understanding of these problems, which may eventually lead to their solution.

ACKNOWLEDGEMENT

I thank Mette Sundblad for bibliographic research and help in formatting this article.

NOTE

1 A possible connection is James M. Sakoda, who apparently was the first person to develop a CA-based model in the social sciences. He published his model in Sakoda 1971, but the basic design of the model was already present in his unpublished dissertation of 1949 (Hegselmann and Flache, 1998: 3.2). However, Schelling has stated to never have heard of Sakoda (Aydinonat, 2005: 5).

REFERENCES

Alexander, S. (1920) *Space, Time, and Deity*. 2 vols. London: Macmillan.

Axtell, R.L., Epstein, J.M., Dean, J.S., Gumerman, G.J., Swedlund, A. C., Harburger, J., Chakravarty, S., Hammond, R., Parker, J. and Parker, M. (2002) 'Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley', *Proceedings of the National Academy of Sciences*, 99(3): 7275–7279.

Aydinonat, N.E. (2005) 'An interview with Thomas C. Schelling: Interpretation of game theory and the checkerboard model', *Economics Bulletin*, 2(2): 1–7.

Aydinonat, N.E. (2008) *The Invisible Hand in Economics: How Economists Explain Unintended Social Consequences, INEM Advances in Economic Methodology*. London: Routledge.

Bankes, S.C., Lempert, R.J. and Popper, S.W. (2001) 'Computer-assisted reasoning', *Computing in Science and Engineering*, 3(2): 71–77.

Bedau, M. (1997) 'Weak emergence', *Philosophical Perspectives*, 11: 375–399.

Berlekamp, E.R., Conway, J.H. and Guy, R.K. (1982) Winning *Ways for your Mathematical Plays*. Volume 2*: Games in Particular*. London: Academic Press.

Boden, M.A. (ed.) (1996) *The Philosophy of Artificial Life. Oxford Readings in Philosophy*. New York: Oxford University Press.

Brouwers, L., Mäkilä, K. and Camitz, M. (2006) 'Spridning av smittkoppor – simuleringsexperiment', *SMI-Rapport*, 5.

Broad, C.D. (1925) *The Mind and Its Place in Nature*. London: Routledge & Kegan Paul.

Bunn, D.W. and Oliveira, F.S. (2003) 'Evaluating individual market power in electricity markets via agent-based simulation', *Annals of Operations Research*, 121: 57–77.

Cederman, L.-E. (2002) 'Endogenizing geopolitical boundaries with agent-based modelling', *Proceedings* *of the National Academy*, 99(suppl. 3): 7796–7303.

Cederman, L.-E. (2005) 'Computational models of social forms: Advancing generative process theory', *American Journal of Sociology*, 110(4): 864–893.

Coops, N.C. and Catling, P.C. (2002) '*Prediction* of the spatial distribution and relative abundance of ground-dwelling mammals using remote sensing imagery and *simulation* models', *Landscape Ecology*, 17(2): 173–188.

D'Arms, J., Batterman, R. and Górny, K. (1998) 'Game theoretic explanations and the evolution of justice', *Philosophy of Science*, 65: 76–102.

Dean, J., Gumerman, G., Epstein, J., Axtell, R., Swedlund, A., Parker, M. and McCarroll, S. (2000) 'Understanding anasazi cultural change through agent-based modelling', in T. Kohler and G. Gumerman (eds) *Dynamics in Human and Primate Societies: Agent-Based Modelling of Social and Spatial Processes.* Santa Fe Institute Studies in the Science of Complexity. New York: Oxford University Press. pp. (179–205)

Doran J.E. (1998) 'Simulating collective misbelief', *Journal of Artificial Societies and Social Simulation*, 1(1): http://www.soc.surrey.ac.uk/JASSS/1/1/3.html.

Dowling, D. (1999) 'Experimenting on theories', *Science in Context*, 12(2): 261–273.

Epstein, J.M. (1999) 'Agent-based computational models and generative social science', *Complexity*, 4(5): 41–57.

Epstein, J. and Axtell, R. (1996) *Growing Artificial Societies: Social Science from the Bottom-Up.* Cambridge, MA: MIT Press.

Eubank, S., Guclu, H., Kumar, V.S.A., Marathe, M., Srinivasan, A., Toroczcai, Z. and Wang, N. (2004) 'Modelling disease outbreaks in realistic urban social networks', *Nature*, 429: 180–184.

Frankfurt, H. (1978) 'The problem of action', *American Philosophical Quarterly*, 15: 157–162.

Frigg, R. and Reiss, J. (2009) 'The philosophy of simulation: Hot new issues or same old stew?', *Synthese*, 169(3): 593–613.

Gilbert, N. and Troitzsch, K.G. (1999) *Simulation for the Social Scientist.* Milton Keynes: Open University Press.

Gilbert, N. (2004) 'Agent-based social simulation: Dealing with complexity', *Centre for Research on Social Simulation Mimeo.* Guildford: University of Surrey.

Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S.F., Thulke, H.-H., Weiner, J., Wiegand, T. and DeAngelis, D.L. (2005) 'Pattern-oriented modelling of agent-based complex systems: Lessons from ecology', *Science*, 310: 987–991.

Grüne-Yanoff, T. (2009) 'The explanatory potential of artificial societies', *Synthese*, 169(3): 539–555.

Grüne-Yanoff, T. (2010) 'Agent-based models as policy decision tools: The case of smallpox vaccination', *Simulation and Gaming.* Published online 7 September 2010. DOI: 10.1177/1046878110377484.

Guala, F. (2005) *The Methodology of Experimental Economics.* Cambridge: Cambridge University Press.

Happe, K., Kellermann, K. and Balmann, A. (2006) 'Agent-based analysis of agricultural policies: An illustration of the agricultural policy simulator AgriPoliS, its adaptation and behaviour', *Ecology and Society*, 11(1): 49.

Haug, E. (1981) 'Engineering safety analysis via destructive numerical experiments', *EUROMECH 12, Polish Academy of Sciences, Engineering Transactions*, 29(1): 39–49.

Hegselmann, R. and Flache, A. (1998) 'Understanding complex social dynamics: A plea for cellular automata based modelling', *Journal of Artificial Societies and Social Simulation*, 1(3): http://www.soc.surrey.ac.uk/JASSS/1/3/1.html.

Hempel, C.G. (1965) *Aspects of Scientific Explanation.* New York: The Free Press.

Huberman, B. and Glance, N. (1993) 'Evolutionary games and computer simulations', *Proceedings of the National Academy of Science*, 90: 7716–7718.

Humphreys, P. (1991) 'Computer simulations', in A. Fine, M. Forbes and L. Wessels (eds) *PSA 1990, 2.* East Lansing: Philosophy of Science Association. pp. 497–506.

Humphreys, P. (2004) *Extending Ourselves: Computational Science, Empiricism, and Scientific Method.* New York: Oxford University Press.

Humphreys, P. (2009) 'The philosophical novelty of computer simulation methods', *Synthese*, 169(3): 615–626.

Küppers, G. and Lenhard, J. (2005) 'Validation of simulation: Patterns in the social and natural sciences', *Journal of Artificial Societies and Social Simulation*, 8(4): http://jasss.soc.surrey.ac.uk/8/4/3.html.

LeBaron, B. (2001) 'Empirical regularities from interacting long and short memory investors in an agent-based financial market', *IEEE Transactions on Evolutionary Computation*, 5: 442–455.

Lehtinen, A. and Kuorikoski, J. (2007) 'Computing the perfect model: Why do economists shun simulation?', *Philosophy of Science*, 74(3): 304–329.

Lempert, R.J. (2002) 'New decision sciences for complex systems', *Proceedings of the National Academy of Sciences*, 99(supplement 3): 7309–7313.

Mäki, U. (1994) 'Isolation, idealization and truth in economics', in B. Hamminga and N. De Marchi (eds) *Idealization-VI: Idealization in Economics. Poznan Studies in the Philosophy of the Sciences and the Humanities*, 38: 7–68. Dordrecht: Kluwer.

Marimon, R., McGrattan, E. and Sargent, T.J. (1990) 'Money as a medium of exchange in an economy with artificially intelligent agents', *Journal of Economic Dynamics and Control*, 14(2): 329–373.

Maynard Smith, J. and Price, G. (1973) 'The logic of animal conflict', *Nature*, 146: 15–18.

Menger, C. (1982) 'On the origin of money', *Economic Journal*, 2: 239–255.

Mill, J.S. (1843) *System of Logic*. London: Longmans, Green, Reader, and Dyer. [8th edn, 1872].

Morgan, M.S. (2003) 'Experiments without material intervention: Model experiments, virtual experiments and virtually experiments', in H. Radder *(ed.) The Philosophy of Scientific Experimentation*. University of Pittsburgh Press. pp. 217–235.

Moss, S. (2002) 'Policy analysis from first principles', *Proceedings of the US National Academy of Sciences*, 99(Suppl. 3): 7267–7274.

Nilsson, N. *(1998) Artificial Intelligence: A New Synthesis*. Morgan Kaufmann Publishers.

O'Connor, T. (1994) 'Emergent properties', *American Philosophical Quarterly*, 31: 91–104.

Pan, X., Han, C.S., Dauber, K.and Law, K.H. (2007) 'A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations', *AI & Society*, 22(2): 113–132.

Parker, W. (2009) 'Does matter really matter? Computer simulations, experiments and materiality', *Synthese*, 169(3): 483–496.

Poole, D., Mackworth, A. and Goebel, R. (1998) *Computational Intelligence: A Logical Approach. New York: Oxford University Press.*

Rademacher, C., Neuert, C., Grundmann, V., Wissel, C. and Grimm, V. (2004) 'Reconstructing spatiotemporal dynamics of Central European natural beech forests: The rule-based forest model BEFORE', *Forest Ecology and Management*, 194: 349–368.

Ray, T.S. (1996) 'An approach to the synthesis of life', in Margaret A. Boden (ed.) *The Philosophy of Artificial Life. Oxford Readings in Philosophy*. New York: Oxford University Press. pp. 111–145.

Reynolds, C. (1987) *'Flocks, herds and schools: A distributed behavioural model', SIGGRAPH '87: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques (Association for Computing Machinery). pp. 25–34.*

Rohrlich, F. (1991) 'Computer simulation in the physical sciences', *Philosophy of Science Association*, 2: 507–518.

Russell, S.J. and Norvig, P. (2003) *Artificial Intelligence: A Modern Approach*. 2nd edn. Upper Saddle River, NJ: Prentice Hall.

Sakoda, J. M. (1971) The checkerboard model of social interaction, *Journal of Mathematical Sociology*, 1: 119–132.

Samanidou, R., Zschischang, E., Stauffer, D. and Lux, T. (2007) 'Agent-based models of financial markets', *Reports on Progress in Physics*, 70: 409–450.

Sawyer, R.K. (2004) 'Social explanation and computational simulation', *Philosophical Explorations*, 7(3): 219–231.

Schelling, T. (1969) 'Models of segregation', *The American Economic Review*, 59(2): 488–493.

Schelling, T. (1971) 'Dynamic models of segregation', *Journal of Mathematical Sociology*, 1: 143–86.

Schweber, S. and Wächter, M. (2000) 'Complex systems, modelling and simulation', *Studies in the History and Philosophy of Modern Physics*, 31(4): 583–609.

Simon, H.A. (1969) *The Sciences of the Artificial*. Cambridge, MA: MIT Press.

Sigmund, K. (2005) 'John Maynard Smith and evolutionary game theory', *Theoretical Population Biology*, 68: 7–10.

Sober, E. (1996) 'Learning from functionalism – Prospects for strong artificial life', in Margaret A. Boden (ed.) *The Philosophy of Artificial Life. Oxford Readings in Philosophy*. New York: Oxford University Press. pp. 361–378.

Stöckler, M. (2000) 'On modelling and simulations as instruments for the study of complex systems', in M. Carrier , G. Massey and L. Ruetsche (eds) *Science at Century's End: Philosophical Questions on the Progress and Limits of Science*. Pittsburgh: University of Pittsburgh Press. pp. 355–373.

Sugden, R. (2000) 'Credible worlds: The status of theoretical models in economics', *Journal of Economic Methodology*, 7(1): 1–31.

Tesfatsion, L. (2006) 'Agent-based computational economics: A constructive approach to economic theory', in L. Tesfatsion and K. Judd (eds) *Handbook of Computational Economics* Vol. 2. North Holland: Elsevier. pp. 1–50.

Winsberg, E. (2001) 'Simulations, models, and theories: Complex physical systems and their representations', *Philosophy of Science*, 68(supplement): S442–S454.

Winsberg, E. (2009) 'A tale of two methods', *Synthese*, 169(3): 575–592.

Wissel, C. (1992) 'Modelling the mosaic-cycle of a Middle European beech forest', *Ecological Modelling*, 63: 29–43.