Agent-Based Computational Models And Generative Social Science

JOSHUA M. EPSTEIN

Economic Studies, Brookings Institution, Washington, D.C., 20036, USA
External Faculty, Santa Fe Institute, Santa Fe, NM, 87501, USA
e-mail: jepstein@brook.edu

This article argues that the agent-based computational model permits a distinctive approach to social science for which the term "generative" is suitable. In defending this terminology, features distinguishing the approach from both "inductive" and "deductive" science are given. Then, the following specific contributions to social science are discussed: The agent-based computational model is a new tool for empirical research. It offers a natural environment for the study of connectionist phenomena in social science. Agent-based modeling provides a powerful way to address certain enduring—and especially interdisciplinary—questions. It allows one to subject certain core theories—such as neoclassical microeconomics—to important types of stress (e.g., the effect of evolving preferences). It permits one to study how rules of individual behavior give rise—or "map up"—to macroscopic regularities and organizations. In turn, one can employ laboratory behavioral research findings to select among competing agent-based ("bottom up") models. The agent-based approach may well have the important effect of decoupling individual rationality from macroscopic equilibrium and of separating decision science from social science more generally. Agent-based modeling offers powerful new forms of hybrid theoretical-computational work; these are particularly relevant to the study of non-equilibrium systems. The agent-based approach invites the interpretation of society as a distributed computational device, and in turn the interpretation of social dynamics as a type of computation. This interpretation raises important foundational issues in social science—some related to intractability, and some to undecidability proper. Finally, since "emergence" figures prominently in this literature, I take up the connection between agent-based modeling and classical emergentism, criticizing the latter and arguing that the two are incompatible. © 1999 John Wiley & Sons, Inc.

1. GENERATIVE SOCIAL SCIENCE

The agent-based computational model—or artificial society—is a new scientific instrument.[1] It can powerfully advance a distinctive approach to social science, one for which the term "generative" seems appropriate. I will discuss this term more fully below, but in a strong form, the central idea is this: To the generativist, explaining the emergence[2] of macroscopic societal regularities, such as norms or price equilibria, requires that one answer the following question:

The Generativist's Question

* *How could the decentralized local interactions of heterogeneous autonomous agents generate the given regularity?*

The agent-based computational model is well-suited to the study of this question since the following features are characteristic[3]:

Heterogeneity

Representative agent methods—common in macroeconomics—are not used in agent-based models.[4] Nor are agents

aggregated into a few homogeneous pools. Rather, agent populations are heterogeneous; individuals may differ in myriad ways—genetically, culturally, by social network, by preferences—all of which may change or adapt endogenously over time.

Autonomy

There is no central, or "top-down," control over individual behavior in agent-based models. Of course, there will generally be feedback from macrostructures to microstructures, as where newborn agents are conditioned by social norms or institutions that have taken shape endogenously through earlier agent interactions. In this sense, micro and macro will typically co-evolve. But as a matter of model specification, no central controllers or other higher authorities are posited *ab initio*.

Explicit Space

Events typically transpire on an explicit space, which may be a landscape of renewable resources, as in Epstein and Axtell (1996), an n-dimensional lattice, or a dynamic social network. The main *desideratum* is that the notion of "local" be well posed.

Local Interactions

Typically, agents interact with neighbors in this space (and perhaps with environmental sites in their vicinity). Uniform mixing is generically not the rule.[5] It is worth noting that—although this next feature is logically distinct from generativity—many computational agent-based models also assume:

Bounded Rationality

There are two components of this: bounded information and bounded computing power. Agents do not have global information, and they do not have infinite computational power. Typically, they make use of simple rules based on local information.[6]

The agent-based model, then, is especially powerful in representing spatially distributed systems of heterogeneous autonomous actors with bounded information and computing capacity who interact locally.

The Generativist's Experiment

In turn, given some macroscopic *explanandum*—a regularity to be explained—the canonical agent-based experiment is as follows:

> ****** *Situate an initial population of autonomous heterogeneous agents in a relevant spatial environment; allow them to interact according to simple local rules, and thereby generate—or "grow"—the macroscopic regularity from the bottom up.*[7]

Concisely, ****** is the way generative social scientists answer *****. In fact, this type of experiment is not new[8] and, in prin-

ciple, it does not necessarily involve computers.[9] However, recent advances in computing, and the advent of large-scale agent-based computational modeling, permit a generative research program to be pursued with unprecedented scope and vigor.

Examples

A range of important social phenomena have been generated in agent-based computational models, including: right-skewed wealth distributions (Epstein and Axtell, 1996), right-skewed firm size and growth rate distributions (Axtell, 1999), price distributions (Bak *et al.,* 1993), spatial settlement patterns (Dean *et al.,* 1999), economic classes (Axtell *et al.,* 1999), price equilibria in decentralized markets (Albin and Foley, 1990; Epstein and Axtell, 1996), trade networks (Tesfatsion, 1995; Epstein and Axtell, 1996), spatial unemployment patterns (Topa, 1997), excess volatility in returns to capital (Bullard and Duffy, 1998), military tactics (Ilachinski, 1997), organizational behaviors (Prietula, Carley, and Gasser, 1998), epidemics (Epstein and Axtell, 1996), traffic congestion patterns (Nagel and Rasmussen, 1994), cultural patterns (Axelrod, 1997; Epstein and Axtell, 1996), alliances (Axelrod, 1993; Cederman, 1997), stock market price time series (Arthur *et al.,* 1997), voting behaviors (Kollman, Miller, and Page, 1992), cooperation in spatial games (Lindgren and Nordahl, 1994; Epstein, 1998; Huberman and Glance, 1993; Nowak and May, 1992; Miller, 1996), and demographic histories (Dean *et al.,* 1999). These examples manifest a wide range of (often implicit) objectives and levels of quantitative testing.

Before discussing specific models, it will be useful to identify certain changes in perspective that this approach may impose on the social sciences. Perhaps the most fundamental of these changes involves explanation itself.

2. EXPLANATION AND GENERATIVE SUFFICIENCY

Agent-based models provide computational demonstrations that a given microspecification is in fact *sufficient to generate* a macrostructure of interest. Agent-based modelers may use statistics to gauge the generative sufficiency of a given microspecification—to test the agreement between real-world and generated macrostructures. (On levels of agreement, see Axtell and Epstein, 1994.) A good fit demonstrates that the target macrostructure—the *explanandum*—be it a wealth distribution, segregation pattern, price equilibrium, norm, or some other macrostructure, is effectively attainable under repeated application of agent-interaction rules: It is *effectively computable by agent-society.* (The view of society as a distributed computational device is developed more fully below). Indeed, *this demonstration is taken as a necessary condition for explanation itself.* To the generativist—concerned with *formation dynamics*—it does not suffice to establish that, if deposited in some macroconfiguration, the system will stay there.

Rather, the generativist wants an account of the configuration's *attainment by a decentralized system of heterogeneous autonomous agents.* Thus, the motto of *generative social science,* if you will, is: If you didn't grow it, you didn't explain its emergence. Or, in the notation of first-order logic:

$$(\forall x)(\neg Gx \supset \neg Ex) \qquad [1]$$

It must be emphasized that the motto applies only to that domain of problems involving the formation or emergence of macroscopic regularities. Proving that some configuration is a Nash equilibrium, for example, arguably *does* explain its persistence, but does not account for its attainment.[10]

Regarding the converse of expression [1], if a microspecification, m, generates a macrostructure of interest, then m is a *candidate* explanation. But it may be a relatively weak candidate; merely generating a macrostructure does not necessarily explain its formation particularly well. Perhaps Barnsley's fern (Barnsley, 1988) is a good mathematical example. The limit object indeed looks very much like a black spleenwort fern. But—under iteration of a certain affine function system—it assembles itself in a completely unbiological way, with the tip first, then a few outer branches, eventually a chunk of root, back to the tip, and so forth—not connectedly from the bottom up (now speaking literally).

It may happen that there are distinct microspecifications having equivalent generative power (their generated macrostructures fit the macro-data equally well). Then, as in any other science, one must do more work, figuring out which of the microspecifications is most tenable empirically. In the context of social science, this may dictate that competing microspecifications with equal generative power be adjudicated experimentally—perhaps in the psychology lab.

In summary, if the microspecification m does not generate the macrostructure x, then m is not a candidate explanation. If m does generate x, it is a candidate.[11] If there is more than one candidate, further work is required at the micro-level to determine which m is the most tenable explanation empirically.[12]

For most of the social sciences, it must be said, the problem of multiple competing generative accounts would be an embarrassment of riches. The immediate agenda is to produce generative accounts *per se.* The principal instrument in this research program is the agent-based computational model. And as the earlier examples suggest, the effort is underway.

This agenda imposes a constructivist (intuitionistic) philosophy on social science.[13] In the air is a foundational debate on the nature of explanation reminiscent of the controversy on foundations of mathematics in the 1920s–30s. Central to that debate was the intuitionists' rejection of nonconstructive existence proofs (see below): their insis-

tence that meaningful "existence in mathematics coincides with constructibility." (Fraenkel and Bar-Hillel, 1958, p. 207). While the specifics are of course different here—and I am not discussing intuitionism in mathematics proper—this is the impulse, the spirit, of the agent-based modelers: If the distributed interactions of heterogeneous agents can't generate it, then we haven't explained its emergence.

Generative Versus Inductive and Deductive

From an epistemological standpoint, generative social science, while empirical (see below), is not *inductive,* at least as that term is typically used in the social sciences (e.g., as where one assembles macroeconomic data and estimates aggregate relations econometrically). (For a nice introduction to general problems of induction, beginning with Hume, see Chalmers, 1982. On inductive logic, see Skyrms, 1986. For Bayesians and their critics, see, respectively, Howson and Urbach, 1993 and Glymour, 1980.)

The relation of generative social science to *deduction* is more subtle. The connection is of particular interest because there is an intellectual tradition in which we account an observation as explained precisely when we can *deduce the proposition expressing that observation from other, more general, propositions.* For example, we *explain* Galileo's leaning Tower of Pisa observation (that heavy and light objects dropped from the same height hit the ground simultaneously) by strictly deducing, from Newton's Second Law and the Law of Universal Gravitation, the following proposition: "The acceleration of a freely falling body near the surface of the earth is independent of its mass." In the present connection, we seek to explain macroscopic social phenomena. And we are requiring that they be generated in an agent-based computational model. Surprisingly, in that event, we can legitimately claim that they are strictly deducible. In particular, if one accepts the Church-Turing thesis, then every computation—including every agent-based computation—can be executed by a suitable register machine (Hodel, 1995; Jeffrey, 1991). It is then a theorem of logic and computability that every program can be simulated by a first-order language. In particular, with N denoting the natural numbers:

Theorem. Let P be a program. There is a first-order language L, and for each $a \in N$ a sentence C(a) of L, such that for all $a \in N$, the P-computation with input a halts ⇔ the sentence C(a) is logically valid.

This theorem allows one to use the recursive unsolvability of the halting problem to establish the recursive unsolvability of the validity problem in first-order logic (see Kleene, 1967). Explicit constructions of the correspondence between register machine programs and the associated logical arguments are laid out in detail by Jeffrey (1991) and Hodel (1995). The point here is that for every computation,

there is a corresponding logical deduction. (And this holds even when the computation involves "stochastic" features, since, on a computer, these are produced by *deterministic* pseudo-random number generation. See Knuth, 1969. Even if one conducts a statistical analysis over some distribution of runs—using different random seeds—each run is itself a deduction. Indeed, it would be quite legitimate to speak, in that case, of a distribution of theorems.[14]) In any case, from a technical standpoint, *generative implies deductive,* a point that will loom large later, when we argue that *agent-based modeling and classical emergentism are incompatible.*

Importantly, however, the converse does not apply: Not all deductive argument has the constructive character of agent-based modeling. Nonconstructive existence proofs are obvious examples. These work as follows: Suppose we wish to prove the existence of an x with some property (e.g., that it is an equilibrium). We take as an axiom the so-called Law of the Excluded Middle that (i) either x exists or x does not exist. Next, we (ii) assume that x does *not* exist, and (iii) derive a contradiction. From this we conclude that (iv) x must exist. But we have failed to exhibit x, or indicate any algorithm that would generate it, patently violating the generative motto [1].[15] The same holds for many nonconstructive proofs in mathematical economics and game theory (e.g., deductions establishing the existence of equilibria using fixed-point theorems). See Lewis (1985). In summary, then, generative implies deductive, but the converse is not true.

Given the differences between agent-based modeling and both inductive and deductive social science, a distinguishing term seems appropriate. The choice of "generative" was inspired by Chomsky's (1965) early usage: Syntactic theory seeks minimal rule systems that are sufficient to generate the structures of interest, grammatical constructions among them.[16] The generated structures of interest here are, of course, social.

Now, at the outset, I claimed that the agent-based computational model was a *scientific* instrument. A fair question, then, is whether agent-based computational modeling offers a powerful new way to do empirical research. I will argue that it does. Interestingly, one of the early efforts involves the seemingly remote fields of archaeology and agent-based computation.

3. EMPIRICAL AGENT-BASED RESEARCH

The Artificial Anasazi project of Dean, Gumerman, Epstein, Axtell, Swedlund, McCarroll, and Parker aims to grow an actual 500-year spatio-temporal demographic history—the population time series and spatial settlement dynamics of the Anasazi—testing against data. The Artificial Anasazi computational model proper is a hybrid in which the physical environment is "real" (reconstructed from dendroclimatalogical and other data) and the agents are artifi-

cial. In particular, we are attempting to model the Kayenta Anasazi of Long House Valley, a small region in northeastern Arizona, over the period 800 to 1300 AD, at which point the Anasazi mysteriously vanished from the Valley. The enigma of the Anasazi has long been a central question in Southwestern archaeology. One basic issue is whether environmental (i.e., subsistence) factors alone can account for their sudden disappearance. Or do other factors—property rights, clan relationships, conflict, disease—have to be admitted to generate the true history? In bringing agents to bear on this controversy, we have the benefits of (a) a very accurate reconstruction of the physical environment (hydrology, aggradation, maize potential, and drought severity) on a square hectare basis for each year of the study period, and (b) an excellent reconstruction of household numbers and locations.

The logic of the exercise has been, first, to digitize the true history—we can now watch it unfold on a digitized map of Longhouse Valley. This data set (what really happened) is the target—the *explanandum.* The aim is to develop, in collaboration with anthropologists, microspecifications—ethnographically plausible rules of agent behavior—that will generate the true history. The computational challenge, in other words, is to place artificial Anasazi where the true ones were in 800 AD and see if—under the postulated rules—the simulated evolution matches the true one. Is the microspecification *empirically adequate,* to use van Fraassen's (1980) phrase?[17] From a contemporary social science standpoint, the research also bears on the adequacy of simple "satisficing" rules—rather than elaborate optimizing ones—to account for the observed behavior.

A comprehensive report on Phase 1 (environmental rules only) of this research is given in Dean *et al.* (1999). The full microspecification, including hypothesized agent rules for choosing residences and farming plots, is elaborated there. The central result is that the purely environmental rules explored thus far account for (retrodict) important features of the Anasazi's demography, including the observed coupling between environmental and population fluctuations, as well as important observed spatial dynamics: agglomerations and zonal occupation series. These rules also generate a precipitous decline in population around 1300. However, they do not generate the outright disappearance that occurred. One interpretation of this finding is that subsistence considerations alone do *not* fully explain the Anasazi's departure, and that institutional or other cultural factors were likely involved. This work thus suggests the power and limits of a purely environmental account, a finding that advances the archaeological debate.

Simply to convey the flavor of these simulations—which unfold as animations on the computer—Figure 1 gives a comparison. Each dot is an Anasazi household. The graphic shows the true situation on the right and a simulation outcome on the left for the year 1144. In both cases, agents are

Date: 1144
Simulation

Date: 1144
History
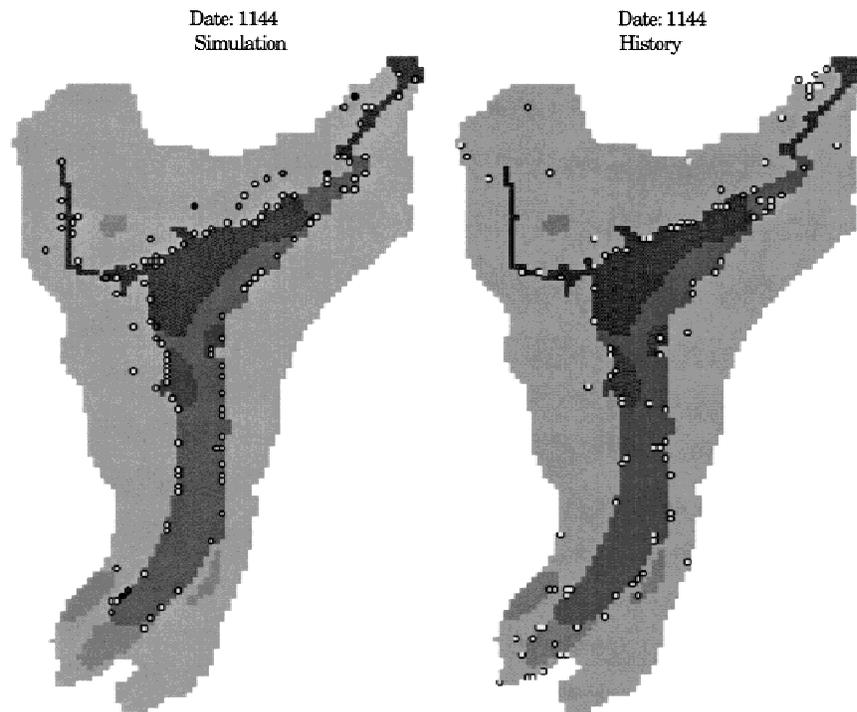
Actual and simulated Anasazi compared
Source: Dean *et al.* (1999, p. 204)

located at the border of the central farming area—associated with a high water table (dark shade)—and the household numbers are interestingly related.

The population time series (see Dean *et al.,* 1999) comparing actual and simulated for a typical run is also revealing. The simulated Anasazi curve is qualitatively encouraging, matching the turning points, including a big crash in 1300, but quantitatively inaccurate, generally overestimating population levels, and failing to generate the "extinction" event of interest.

As noted earlier, one intriguing interpretation of these results is that the Valley could have supported the Anasazi in fact, so their departure may have been the result of *institutional* factors not captured in the purely environmental account.

The claim is not that the current model has solved—or that the planned extensions will ultimately solve—the mystery of the Anasazi. Rather, the point is that agent-based modeling permits a new kind of empirical research (and, it might be noted, a novel kind of interdisciplinary collaboration).

This is by no means the only example of data-driven empirical research with agents. For example, Axtell (1999)

gives an agent-based computational model of firm formation that generates distributions of firm sizes and growth rates close to those observed in the U.S. economy. Specifically, citing the work of Stanley *et al.* (1996, p. 806), Axtell writes that "there are three important empirical facts that an *accurate* theory of the firm should reproduce: (a) firm sizes must be right-skewed, approximating a power law; (b) firm growth rates must be Laplace distributed; (c) the standard deviation in log growth rates as a function of size must follow a power law with exponent -0.15 ± 0.03." He further requires that the model be written at the level of individual human agents—that it be methodologically individualist. Aside from his own agent-based computational model, Axtell writes, " . . . theories of the firm that satisfy all these requirements are unknown to us." (Axtell, 1999, p. 88).

Similarly, observed empirical size-frequency distributions for traffic jams are generated in the agent-based model of Nagel and Rasmussen (1994). Bak, Paczuski, and Shubik (1996) present an agent-based trading model that succeeds in generating the relevant statistical distribution of prices. Axelrod (1993) develops an agent-based model of alliance formation that generates the alignment of seventeen nations in the Second World War with high fidelity. Other exercises in which agent-based models are confronted with data include Kirman and Vriend (1998) and Arthur *et al.* (1997).

As in the case of the Anasazi work, I am not claiming that any of these models permanently resolves the empirical question it addresses. The claim, rather, is that agent-based modeling is a powerful empirical technique. In some of these cases (e.g., Axtell, 1999), the agents are individual humans, and in others (Dean *et al.,* 1999; Axelrod, 1993) they are not. But, in all these cases, the empirical issue is the same: *Does the hypothesized microspecification suffice to generate the observed phenomenon*—be it a stationary firm size distribution, a pattern of alliances, or a nonequilibrium price time series? The answer may be yes and, crucially, it may be no. Indeed, it is precisely the latter possibility—

empirical falsifiability—that qualifies the agent-based computational model as a *scientific* instrument.

In addition to "hard" quantitative empirical targets, agent-based computational models may aim to generate important social phenomena qualitatively. Examples of "stylized facts" generated in such models include: right-skewed wealth distributions (Epstein and Axtell, 1996), cultural differentiation (Epstein and Axtell, 1996; Axelrod 1997), multipolarity in interstate systems (Cederman, 1997), new political actors (Axelrod, 1995), epidemics (Epstein and Axtell, 1996), economic classes (Axtell, Epstein, and Young, 1999), and the dynamics of retirement (Axtell and Epstein, 1999) to name a few. This "computational theorizing,"[18] if you will, can offer basic insights of the sort exemplified in Schelling's (1971) pioneering models of racial segregation, and may, of course, evolve into models directly comparable to data. Indeed, they may inspire the collection of data not yet in hand. (Without theory, it is not always clear what data to collect.) Turning from empirical phenomena, the generated phenomenon may be computation itself.

4. CONNECTIONIST SOCIAL SCIENCE

Certain social systems, such as trade networks (markets), are essentially computational architectures. They are distributed, asynchronous, and decentralized and have endogenous dynamic connection topologies. For example, the CD-ROM version of Epstein and Axtell (1996) presents animations of dynamic endogenous trade networks. (For other work on endogenous trade networks, see Tesfatsion, 1995.) There, agents are represented as nodes, and lines joining agents represent trades. The connection pattern—computing architecture—changes as agents move about and interact economically, as shown in Figure 2.

Whether they realize it or not, when economists say "the market arrives at equilibrium," they are asserting that this type of dynamic "social neural net" has executed a computation—it has computed P*, an equilibrium price vector. No individual has tried to compute this, but the society of agents does so nonetheless. Similarly, convergence to social norms, convergence to strategy distributions (in n-person games), or convergence to stable cultural or even settlement patterns (as in the Anasazi case) are all *social computations* in this sense.

It is clear that the efficiency—indeed the very feasibility—of a social computation may depend on the way in which agents are connected. After all, information in society is not manna from heaven; it is collected and processed at the agent level and transmitted through interaction structures that are endogenous. How then does the endogenous connectivity—the topology—of a social network affect its performance as a distributed computational device, one that, for example, computes price equilibria, or converges to (computes) social norms, or converges to spatial settlement patterns such as cities?[19] Agent-based models allow us to pursue such connectionist social science questions in new and systematic ways.

FIGURE 2
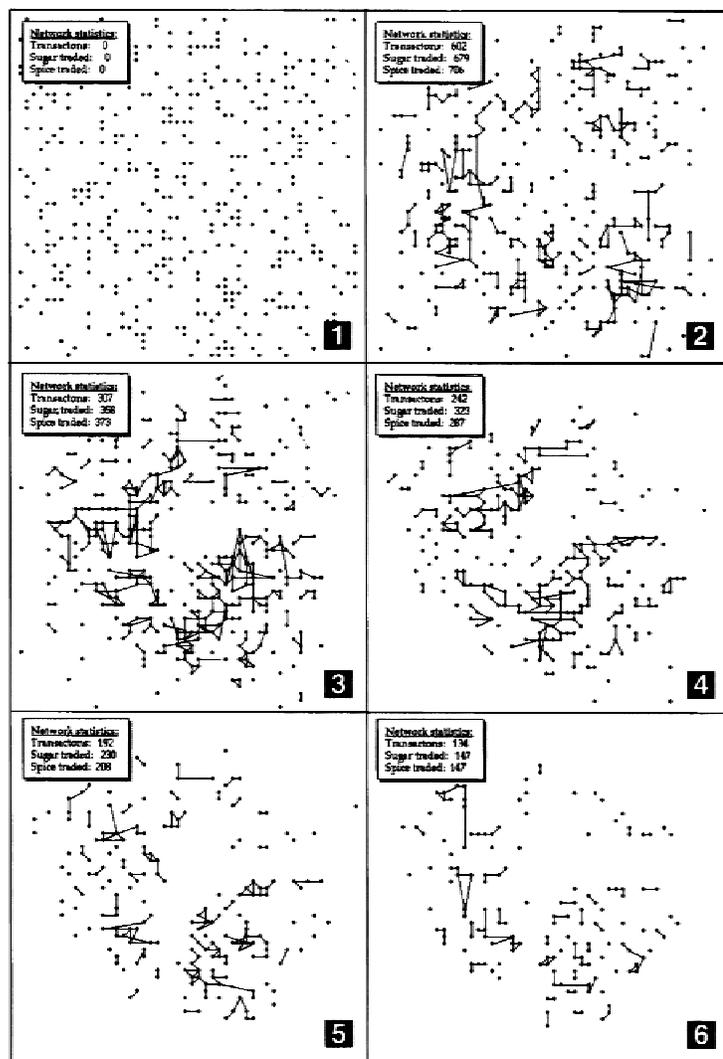
Endogenous trade network
Source: Epstein and Axtell (1996, p. 132)

5. INTERDISCIPLINARY SOCIAL SCIENCE

Many important social processes are not neatly decomposable into separate subprocesses—economic, demographic, cultural, spatial—whose isolated analysis can be somehow "aggregated" to yield an adequate analysis of the process as whole. Yet this is exactly how academic social science is organized—into more or less insular departments and journals of economics, demography, anthropology, and so on. While many social scientists would agree that these divisions are artificial, they would argue that there is no "natural methodology" for studying these processes together, as they interact, though attempts have been made. Social scientists have taken highly aggregated mathematical models—of entire national economies, political systems, and so on—and have "connected" them, yielding "mega-models" that have been attacked on several grounds (see Nordhaus,1992). But attacks on specific models have had the effect of discrediting interdisciplinary inquiry itself, and this is most unfortunate. The line of inquiry remains crucially important. And agent-based modeling offers an alternative, and very natural, technique.

For example, in the agent-based model Sugarscape (Epstein and Axtell, 1996), each individual agent has simple local rules governing movement, sexual reproduction, trading behavior, combat, interaction with the environment, and the transmission of cultural attributes and diseases. These rules can all be "active" at once. When an initial population of such agents is released into an artificial environment in which, and with which, they interact, the resulting artificial society unavoidably links demography, economics, cultural adaptation, genetic evolution, combat, environmental effects, and epidemiology. *Because the individual is multi-dimensional, so is the society.*

Now, obviously, not all social phenomena involve such diverse spheres of life. If one is interested in modeling short-term price dynamics in a local fish market, then human immune learning and epidemic processes may not be relevant. But if one wishes to capture long-term social dynamics of the sort discussed in William McNeill's 1976 book *Plagues and Peoples,* they are essential. Agent-based modelers do not insist that everything be studied all at once. The claim is that the new techniques *allow* us to transcend certain artificial boundaries that may limit our insight.

Nature–Nurture

For example, Sugarscape agents (Epstein and Axtell, 1996) engage in sexual reproduction, transmitting genes for, *inter alia,* vision (the distance they can see in foraging for sugar). An offspring's vision is determined by strictly Mendelian (one locus-two allele) genetics, with equal probability of inheriting the father's or mother's vision. One can easily plot average vision in society over time. Selection will favor agents with relatively high vision—since they'll do better in the competition to find sugar—and, as good Darwinians, we expect to see average vision increase over time, which it does. Now, suppose we wish to study the effect of various *social* conventions on this *biological* evolution. What, for example, is the effect of inheritance—the social convention of passing on accumulated sugar wealth to offspring—on the curve of average vision? Neither traditional economics nor traditional population genetics offer particularly natural ways to study this sort of "nature–nurture" problem. But they are naturally studied in an agent-based artificial society: Just turn inheritance "off" in one run and "on" in another, and compare![20] Figure 3 gives a typical realization.

With inheritance, the average vision curve (gray) is lower: Inheritance "dilutes" selection. Because they inherit sugar, the offspring of wealthy agents are buffered from selection pressure. Hence, low-vision genes persist that would be selected out in the absence of this social convention. We do not offer this is a general law, nor are we claiming that agent-based models are the only ones permitting exploration of such topics.[21] The claim is that they offer a new, and particularly natural, methodology for approaching certain interdisciplinary questions, including this one. Some of

FIGURE 3
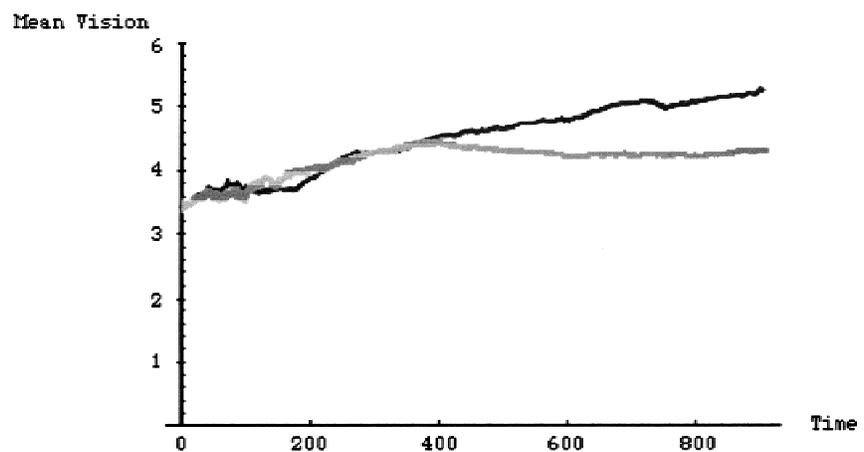
Effect of inheritance on selection
Source: Epstein and Axtell (1996, p. 68)

these questions can be posed in ways that subject dominant theories to stress.

6. THEORY STRESSING

One can use agent-based models to test the robustness of standard theory. Specifically, one can relax assumptions about individual—micro level—behavior and see if standard results on macro behavior collapse. For example, in neoclassical microeconomic theory, individual preferences are assumed to be fixed for the lifetime of the agent. On this assumption (and certain others), individual utility maximization leads to price equilibrium and allocative efficiency (the First Welfare Theorem). But, what if individual preferences are *not* fixed but vary culturally? In Epstein and Axtell (1996), we introduce this assumption into trading agents who are neoclassical in all other respects (e.g., they have Cobb-Douglas utility functions and engage only in Pareto-improving trades with neighbors). The result is far-from-equilibrium markets. The standard theory is not robust to this relaxation in a core assumption about individual behavior. For a review of the literature on this central fixed preferences assumption, see Bowles (1998).

7. AGENTS, BEHAVIORAL SOCIAL SCIENCE, AND THE MICRO-MACRO MAPPING

What can agent-based modeling and behavioral research do for one another? It is hard to pinpoint the dawn of experimental economics, though Simon (1996) credits Katona (1951) with the fundamental studies of expectation formation. In any event, there has been a resurgence of important laboratory and other experimental work on individual decisionmaking over the last two decades. See, for example, Camerer (1997), Rabin (1998), Camerer and Thaler (1995), Tversky and Kahneman (1986), and Kagel and Roth (1995). This body of laboratory social science, if I may call it that, is giving us an ever-clearer picture of how *homo sapiens*—as against *homo economicus*—actually makes decisions. However, a crucial lesson of Schelling's segregation model, and of many subsequent Cellular Automaton models, such as "Life"—not to mention agent-based models themselves—is that *even perfect knowledge of individual decision rules does not always allow us to predict macroscopic structure.* We get macro-surprises despite complete micro-knowledge. Agent-based models allow us to study the micro-to-macro mapping. It is obviously essential to begin with solid foundations regarding individual behavior, and behavioral research is closing in on these. However, we will still need techniques for "projecting up" to the macro level from there (particularly for spatially-distributed systems of heterogeneous individuals). Agent modeling offers behavioral social science a powerful way to do that.

Agent-based models may also furnish laboratory research with counterintuitive hypotheses regarding individual behavior. Some, apparently bizarre, system of individual agent rules may generate macrostructures that mimic the observed ones. Is it possible that those are, in fact, the operative micro-rules? It might be fruitful to design laboratory experiments to test hypotheses arising from the unexpected generative sufficiency of certain rules.

What does behavioral research offer agent-based modeling? Earlier, we noted that different agent-based models might have equal generative (explanatory) power and that, in such cases, further work would be necessary to adjudicate between them. But if two models are doing equally well in generating the macrostructure, preference should go to the one that is best at the micro level. So, if we took the two microspecifications as competing hypotheses about individual behavior, then—apropos of the preceding remark—behavioral experiments might be designed to identify the better hypothesis (microspecification) and, in turn, the better agent model. These, then, are further ways in which agent-based computational modeling can contribute to empirical social science research.

8. DECOUPLINGS

As noted earlier, to adopt agent-based modeling does not compel one to adopt methodological individualism. However, extreme methodological individualism is certainly possible (indeed common) in agent-based models. And individual-based models may have the important effect of *decoupling individual rationality from macroscopic equilibrium.* For example, in the individual-based retirement model of Axtell and Epstein (1999), macroscopic equilibrium is attained—through a process of imitation in social networks—even though the vast preponderance of individuals are *not* rational. Hence—as in much evolutionary modeling—micro rationality is not a *necessary* condition for the attainment of macro equilibrium.[22] Now, we also have agent-models in which macroscopic equilibrium is not attained despite orthodox utility maximization at the individual level. The non-equilibrium economy under evolving preferences (Epstein and Axtell, 1996) noted earlier is an example. Hence, micro rationality is not a *sufficient* condition for macro equilibrium. But if individual rationality is thus neither necessary nor sufficient for macro equilibrium, the two are logically independent—or decoupled, if you will.

Now, the *fraction* of agents in an imitative system (such as the retirement model) who are rational will definitely affect the *rate* at which any selected equilibrium sets in. But the asymptotic equilibrium behavior *per se* does not depend on the dial of rationality, despite much behavioral research on this latter topic. Perhaps the main issue is not how much rationality there is (at the micro level), but *how little is enough* to generate the macro equilibrium.

In passing, it is worth noting that this is of course a huge issue for policy, where "fad creation" may be far more ef-

fective than real education. Often, the aim is not to equip target populations with the data and analytical tools needed to make rational choices; rather, one displays exemplars and then presses for mindless imitation. "Just say no to drugs" not because it's rational—in a calculus of expected lifetime earnings—but because a famous athlete says "no" and it's a norm to imitate him. The manipulation of uncritical imitative impulses may be more effective in getting to a desired macro equilibrium than policies based on individual rationality. The social problem, of course, is that populations of uncritical imitators are also easy fodder for lynch mobs, witch hunts, Nazi parties, and so forth. Agent-based modeling is certainly not the only way to study social contagion (see, for example, Kuran, 1989), but it is a particularly powerful way when the phenomenon is spatial and the population in question is heterogeneous.

Relatedly, agent-based approaches may decouple social science from decision science. In the main, individuals do not *decide*—they do not *choose*—in any sensible meaning of that term, to be ethnic Serbs, to be native Spanish speakers, or to consider monkey brain a delicacy. Game theory may do an interesting job explaining the decision of one ethnic group to attack another at a certain place or time, but it doesn't explain how the ethnic group arises in the first place or how the ethnic divisions are transmitted across the generations. Similarly for economics, what makes monkey brain a delicacy in one society and not in another? Cultural (including preference) patterns, and their nonlinear tippings are topics of study in their own right with agents.[23] See Axelrod (1997) and Epstein and Axtell (1996).

9. ANALYTICAL-COMPUTATIONAL APPROACH TO NON-EQUILIBRIUM SOCIAL SYSTEMS

For many social systems, it is possible to prove deep theorems about asymptotic equilibria. However, the time required for the system to attain (or closely approximate) such equilibria can be astronomical. The transient, out-of-equilibrium dynamics of the system are then of fundamental interest. A powerful approach is to combine analytical proofs regarding asymptotic equilibria with agent-based computational analyses of long-lived transient behaviors, the meta-stability of certain attractors, and broken ergodicity in social systems.

One example of this hybrid analytical-computational approach is Axtell, Young, and Epstein (1999). We develop an agent-based model to study the emergence and stability of equity norms in society. (In that article, we explicitly define the term "emergent" to mean simply "arising from decentralized bilateral agent-interactions.") Specifically, agents with finite memory play Best Reply to Recent Sample Evidence (Young, 1995, 1998) in a three-strategy Nash Demand Game, and condition on an arbitrary "tag" (e.g., a color) that initially has no social or economic significance—it is simply a distinguishing mark. Expectations are generated endogenously through bilateral interactions. And, over time, these tags acquire socially organizing salience. In particular, tag-based classes arise. (The phenomenon is akin to the evolution of meaning discussed in Skyrms, 1998.)

Now, introducing noise, it is possible to cast the entire model as a Markov process and to prove rigorously that it has a unique stationary strategy distribution. When the noise level is positive and sufficiently small, the following asymptotic result can be proved: *The state with the highest long-run probability is the equity norm, both between and within groups.*

Salutary as this asymptotic result may appear, the transition from inequitable states to these equitable ones can be subject to tremendous inertia. Agent-based models allow us to systematically study long-lived transient behaviors. We know that, beginning in an inequitable regime, the system will ultimately "tip" into the equity norm. But how does the waiting time to this transition depend on the number of agents and on memory length? In this case, the waiting time scales *exponentially in memory length, m, and exponentially in N, the number of agents.* Overall, then, the waiting time is immense for m=10 and merely N=100, for example.

Speaking rigorously, the equity norm is stochastically stable (see Young, 1998). The agent-based computational model reveals, however, that—depending on the number of agents and their memory lengths—the waiting time to transit from an inequitable regime to the equitable one may be astronomically long.

This combination of formal (asymptotic) and agent-based (non-equilibrium) analysis seems to offer insights unavailable from either approach alone, and to represent a useful hybrid form of analytical-computational study. For sophisticated work relating individual-based models to analytical ones in biology, see Flierl, et al. (1999).

10. FOUNDATIONAL ISSUES

We noted earlier that markets can be seen as massively parallel spatially distributed computational devices with agents as processing nodes. To say that "the market clears" is to say that this device has completed a computation. Similarly, convergence to social norms, convergence to strategy distributions (in n-person games), or convergence to stable cultural or settlement patterns, are all social computations in this sense. Minsky's (1985) famous phrase was "the Society *of* Mind." What I'm interested in here is "the Society *as* Mind," society as a computational device. (On that strain of functionalism that would be involved in literally asserting that a society could be a mind, see Sober, 1996.)

Now, once we say "computation" we think of Turing machines (or, equivalently, of partial recursive functions). In the context of n-person games, for example, the isomorphism with societies is direct: Initial strategies are tallies on

a Turing machine's input tape; agent interactions function to update the strategies (tallies) and thus represent the machine's state transition function; an equilibrium is a halting state of the machine; the equilibrium strategy distribution is given by the tape contents in the halting state; and initial strategy distributions that run to equilibrium are languages accepted by the machine. The isomorphism is clear. Now, we know what makes for an intractable, or "hard," computational problem. So, given our isomorphism, is there a computational answer to the question, "What's a hard social problem?"

A Computational Characterization of Hard Social Problems

In the model of tag-based classes discussed earlier (Axtell *et al.,* 1999), we prove rigorously that, asymptotically, the equity norm will set in. However, beginning from any other (meta-stable) equilibrium, the time to transit into the equitable state scales exponentially in the number of agents and exponentially in the agents' memory length. If we adopt the definition that social states are *hard to attain* if they are *not effectively computable by agent society in polynomial time,* then equity is hard. (The point applies to this particular set-up; I am emphatically not claiming that there is anything immutable about social inequity.) In a number of models, the analogous point applies to economic equilibria: There are nonconstructive proofs of their existence but computational arguments that their *attainment* requires time that scales exponentially in, for instance, the dimension of the commodity space.[24] On our tentative definition, then, computation of (attainment of) economic equilibria would qualify as another *hard social problem.*

So far we have been concerned with the question, "Does an initial social state run to equilibrium?" or, equivalently, "Does the machine halt given input tape x?" Now, like satisfiability, or truth-table validity in sentential logic, these problems are in principle decidable (that is, the equilibria are effectively computable), but not on time scales of interest to humans. [Here, with Simon (1978), we use the term "time" to denote "the number of elementary computation steps that must be executed to solve the problem."]

Godelian Limits

But there are social science problems that are undecidable in principle, now in the sense of Godel or the Halting Problem. Rabin (1957) showed that "there are actual win-lose games which are strictly determined for which there is no effectively computable winning strategy." He continues, "Intuitively, our result means that there are games in which the player who in theory can always win, cannot do so in practice because it is impossible to supply him with effective instructions regarding how he should play in order to win." Another nice example, based on the unsolvability of Hilbert's Tenth Problem, is given by Prasad (1991):

For n-player games with polynomial utility functions and natural number strategy sets the problem of finding an equilibrium is not computable. There does not exist an algorithm which will decide, for any such game, whether it has an equilibrium or not . . . When the class of games is specified by a finite set of players, whose choice sets are natural numbers, and payoffs are given by polynomial functions, the problem of devising a procedure which computes Nash equilibria is unsolvable.

Other results of comparable strength have been obtained by Lewis (1985,1992a, and 1992b).[25]

Implications for Rational Choice Theory

Here lies the deepest conceivable critique of rational choice theory. There are strategic settings in which the individually optimizing behavior is uncomputable in principle. A second powerful critique is that, while possible in principle, optimization is computationally intractable. As Duncan Foley summarizes, "The theory of computability and computational complexity suggest that there are two inherent limitations to the rational choice paradigm. One limitation stems from the possibility that the agent's problem is in fact undecidable, so that no computational procedure exists which for all inputs will give her the needed answer in finite time. A second limitation is posed by computational complexity in that even if her problem is decidable, the computational cost of solving it may in many situations be so large as to overwhelm any possible gains from the optimal choice of action." (See Albin, 1998; p. 46.) For a fundamental statement, see Simon (1978).

These possibilities are disturbing to many economists. They implicitly believe that if the individual is insufficiently rational it must follow that decentralized behavior is doomed to produce suboptimality at the aggregate level. The invisible hand requires rational fingers, if you will. There are doubtless cases in which this holds. But it is not so in all cases. As noted earlier, in the retirement model of Axtell and Epstein (1999), as well as in much evolutionary modeling, an ensemble of locally interacting agents—none of whom are canonically rational—can nonetheless attain efficiency in the aggregate. Even here, of course, issues of exponential waiting time arise (as in the classes model above). But it is important to sort the issues out.

The agent-based approach forces on us the interpretation of society as a computational device, and this immediately raises foundational specters of computational intractability and undecidability. Much of the economic complexity literature concerns the uncomputability of optimal strategies by *individual rational agents,* surely an important issue. However, our central concern is with the effective computability (attainment) of equilibria *by societies of boundedly rational agents.* In that case, it is irrelevant that

equilibrium can be computed by an economist external to the system using the Scarf, or other such, algorithm. The entire issue is whether it can be attained—generated—through decentralized local interactions of heterogeneous boundedly rational actors. And the agent-based computational model is a powerful tool in exploring that central issue. In some settings, it may be the only tool.

11. EQUATIONS VERSUS AGENT-BASED MODELS

Three questions arise frequently and deserve treatment: Given an agent-based model, are there equivalent equations? Can one "understand" one's computational model without such equations? If one has equations for the macroscopic regularities, why does one need the "bottom-up" agent model?

Regarding the first question—are there equivalent equations for every computational model—the answer is immediate and unequivocal: absolutely. On the Church-Turing Thesis, every computation (and hence every agent-based model) can be implemented by a Turing machine. For every Turing machine there is a unique corresponding and equivalent Partial Recursive Function (see Rogers, 1967). Hence, in principle, for any computation there exist equivalent equations (involving recursive functions). Alternatively, any computer model uses some finite set of memory locations, which are updated as the program executes. One can think of each location as a variable in a discrete dynamical system. In principle, there is some—perhaps very high dimensional—set of equations describing those discrete dynamics. Now, could a human write the equations out? Solve them or even find their equilibria (if such exist)? The answer is not clear. If the equations are meant to represent large populations of discrete heterogeneous agents coevolving on a separate space, with which they interact, it is not obvious how to formulate the equations, or how to solve them if formulated. And, for certain classes of problems (e.g., the PSPACE Complete problems), it can be proved rigorously that simulation is—in a definite sense—the best one can do in principle (see Buss *et al.,* 1991). But that does not mean—turning to the second question—we have no idea what's going on in the model.

To be sure, a theorem is better than no theorem. And many complex social phenomena may ultimately yield to analytical methods of the sort being pioneered by Young (1998), Durlauf (1997), and others. But

an experimental attitude is also appropriate. Consider biology. No one would fault a "theoremless" laboratory biologist for claiming to understand population dynamics in beetles when he reports a regularity observed over a large number of experiments. But when agent-based modelers show such results—indeed, far more robust ones—there's a demand for equations and proofs. These would be valuable, and we should endeavor to produce them. Meanwhile, one can do perfectly legitimate "laboratory" science with computers, sweeping the parameter space of one's model, and conducting extensive sensitivity analysis, and claiming substantial understanding of the relationships between model inputs and model outputs, just as in any other empirical science for which general laws are not yet in hand.[26]

The third question involves confusion between explanation and description, and might best be addressed through an example. In Epstein and Axtell (1996), spatially distributed local agent interactions generate the oscillatory aggregate population time series shown in Figure 4.

The question then arises: Could you not get that same curve from some low-dimensional differential equation, and if so, why do you need the agent model? Let us imagine that we can formulate and analytically solve such an equation, and that the population trajectory is exactly $P(t) = A+B \sin(Ct)$ for constants A, B, and C. Now, *what is the explanatory significance of that descriptively accurate result?* It depends on one's criteria for explanation. If we are generativists, the question is: How could the spatially decentralized interactions of heterogeneous autonomous agents generate that macroscopic regularity? If that is one's question, then the mere formula $P(t) = A+B \sin(Ct)$ is *devoid of explanatory power despite its descriptive accuracy.* The choice of agents versus equations always hinges on the objectives of the

FIGURE 4
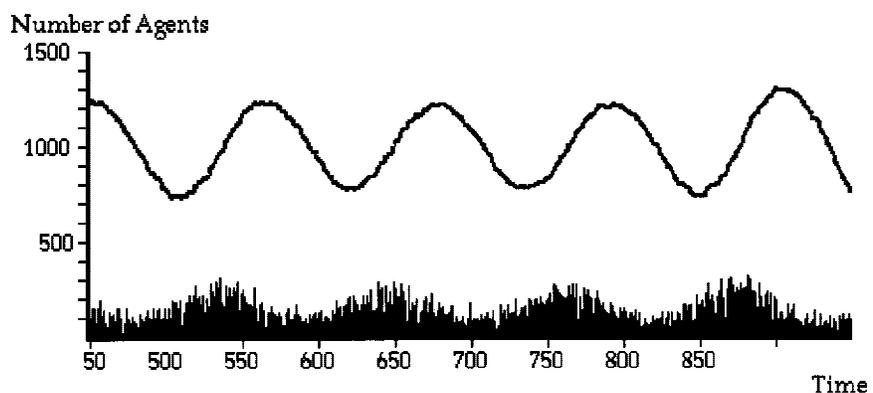
Oscillatory population time series
Source: Epstein and Axtell (1996, p. 161); black vertical spikes represent births

analysis. Given some perfectly legitimate objectives, differential equations are the tool of choice; given others, they're not. If we are explicit as to our objectives, or explanatory criteria, no confusion need arise. And it may be that hybrid models of a second sort are obtainable in which the macrodynamics are well described by an explicit low-dimensional mathematical model, but are also generated from the bottom up in a model population of heterogeneous autonomous agents. That would be a powerful combination. In addition to important opportunities, the field of agent-based modeling, like any young discipline, faces a number of challenges.

12. CHALLENGES

First, the field lacks standards for model comparison and replication of results; see Axtell *et al.* (1996). Implicit in this is the need for standards in reportage of assumptions and certain procedures. Subtle differences can have momentous consequences. For example, how, exactly, are agents being updated? The Huberman and Glance (1993) critique of Nowak and May (1992) is striking proof that asynchronous updating of agents produces radically different results from synchronous updating. Huberman and Glance show that Nowak and May's main result—the persistence of cooperation in a spatial Prisoner's Dilemma game—depends crucially on synchronous updating. When, *ceteris paribus,* Huberman and Glance introduce asynchronous updating into the Nowak and May model, the result is convergence to pure defection. (For a spatial Prisoner's Dilemma model with asynchronous updating in which cooperation *can* persist, see Epstein, 1998.) The same sorts of issues arise in randomizing the agent call order, where various methods— with different effects on output—are possible.

It is also fair to say that solution concepts are weak. Certainly, hitting the "Go" button and watching the screen does not qualify as *solving* anything—any more than an evening at the casino *solves* the Gambler's Ruin Problem from Markov Theory. An individual model run offers a sample path of a (typically) stochastic process, but that is not a general solution—a specific element of some well-defined function space (e.g., a Hilbert or Sobolev space). As noted earlier, it is often possible to sweep the parameter space of one's model quite systematically and thereby obtain a statistical portrait of the relationship between inputs and outputs, as in Axelrod (1997) or Epstein (1998). But it is fair to say that this practice has not been institutionalized.

A deeper issue is that sweeping a model's numerical parameter space is easier than exploring the space of possible agent behavioral *rules* (e.g., "If a neighboring agent is bigger than you, run away" or "Always share food with kin agents"). For artificial societies of any complexity (e.g., Sugarscape), we have no efficient method of searching the space of possible individual rules for those that exhibit generative power. One can imagine using evolutionary approaches to this. First, one would define a metric such that, given a microspecification, the distance from model outputs (generated macrostructures) to targets (observed macrostructures) could be computed. The better the match (the smaller this distance) the "fitter" is the microspecification. Second, one would encode the space of candidate microspecifications and turn, say, a Genetic Algorithm (GA) (see Holland, 1992; Mitchell, 1998) loose on it. The GA might turn up counterintuitive *boundedly* rational rules that are highly "fit" in this sense of generating macrostructures "close" to the targets. (These then become hypotheses for behavioral research, as discussed earlier.)

This strikes me as a far more useful application of GAs than the usual one: finding *hyper-rational* individual strategies, which we now have strong experimental evidence are *not* being employed by humans. The problem is how to encode the vast space of possible individual rules (not to mention the raw computational challenge of searching it once encoded). In some restricted cases, this has been done successfully (Axelrod, 1987; Crutchfield and Mitchell, 1995), but for high dimensional agents engaged in myriad social interactions—economic, cultural, demographic—it is far from clear how to proceed.

One of the central concepts in dynamics is sensitivity. Sensitivity involves the effect on output (generated macrostructure) of small changes in input (microspecification). To assess sensitivity in agent models, we have to do more than encode the space of rules—we have to *metrize* it. To clarify the issue, consider the following agent rules (*methods* of agent-objects):

- Rule a = Never attack neighbors.
- Rule b = Attack a neighbor if he's green.
- Rule c = Attack a neighbor if he's smaller than you.

Which rule—b or c—represents a "smaller departure from" Rule a? Obviously, the question is ill-posed. And yet we speak of "small changes in the rules" of agent-based models. Some areas (e.g., Cellular Automata) admit binary encodings of rule space for which certain metrics—taxicab or Hamming distance—are natural. But for artificial societies generally, no such simple avenues present themselves. What then constitutes a *small* rule change? Without some metric, we really cannot develop the analogue, for agent-based models, of structural stability—or, equivalently, of bifurcation theory—in dynamical systems.

Some challenges are sociological. Generating collective behavior that to the naked eye "looks like flocking" can be extremely valuable, but it is a radically different enterprise from generating, say, a specific distribution of wealth with parameters close to those observed in society. Crude qualitative caricature is a perfectly respectable goal. But if that is one's goal, the fact must be stated explicitly—perhaps using

the terminology proposed in Axtell and Epstein (1994). This will avert needless resistance from other fields where "normal science" proceeds under established empirical standards patently not met by cartoon "boid" flocks, however stimulating and pedagogically valuable these may be. On the pedagogical value of agent-based simulation generally, see Resnick (1994).

A number of other challenges include building community and sharing results and are covered in Axelrod (1997c). In addition to foundational, procedural, and other scientific challenges, the field of "complexity" and agent-based modeling faces terminological ones. In particular, the term "emergence" figures very prominently in this literature. It warrants an audit.

13. "EMERGENCE"

I have always been uncomfortable with the vagueness—and occasional mysticism—surrounding this word and, accordingly, tried to define it quite narrowly in Epstein and Axtell (1996). There, we defined "emergent phenomena" to be simply "stable macroscopic patterns arising from local interaction of agents."[27] Many researchers define the term in the same straightforward way (e.g., Axelrod, 1997c). Since our work's publication, I have researched this term more deeply and find myself questioning its adoption altogether.

"Emergence" has a history, and it is an extremely spotty one, beginning with classical British emergentism in the 1920s and the works of Samuel Alexander (*Space, Time, and Deity,* 1920), C.D. Broad (*The Mind and its Place in Nature,* 1925), and C. Lloyd Morgan (*Emergent Evolution,* 1923). The complexity community should be alerted to this history. There is an unmistakably anti-scientific—even deistic—flavor to this movement, which claimed *absolute unexplainability* for emergent phenomena. In the view of these authors, emergent phenomena are unexplainable in principle. "The existence of emergent qualities . . . admits no explanation," wrote Alexander (1920).[28] As philosopher Terence Horgan recounts, emergent phenomena were to be "accepted (in Samuel Alexander's striking phrase) 'with natural piety'."

Striking indeed, this sort of language, and classical emergentism's avowedly vitalist cast (see Morgan, 1923) stimulated a vigorous—and to my mind, annihilative—attack by philosophers of science. In particular, Hempel and Oppenheim (1948) wrote, "This version of emergence . . . is objectionable not only because it involves and perpetuates certain logical confusions but also because not unlike the ideas of neovitalism, it encourages an attitude of resignation which is stifling to scientific research. No doubt it is this characteristic, together with its theoretical sterility, which accounts for the rejection, by the majority of contemporary scientists, of the classical absolutist doctrine of emergence."

Classical absolute emergentism is encapsulated nicely in the following formalization of Broad's (1925, p. 61):

Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other . . . and that the characteristic properties of the whole R(A,B,C) cannot, even in theory, be *deduced* from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not in the form R(A,B,C) [emphasis in original].

Agent-Based Modeling Incompatible with Classical Emergentism

Before explicating the logical confusion noted by Hempel and Oppenheim, we can fruitfully apply a bit of logic ourselves. Notice that we have actually accumulated a number of first-order propositions. For predicates, let C stand for classically emergent, D for deducible, E for explained, and G for generated (in a computational model). Then, if x is a system property, we have:

(1) $(\forall x)(Cx \supset \neg Dx)$ Broad (emergent implies not deducible)[29]

(2) $(\forall x)(Cx \supset \neg Ex)$ Alexander (emergent implies not explainable)

(3) $(\forall x)(\neg Gx \supset \neg Ex)$ Generativist Motto (not generated implies not explained)

(4) $(\forall x)(Gx \supset Dx)$ Theorem (generated implies deduced)

Although a number of derivations are possible,[30] the essential point involves (1) and (4). By the earlier Theorem (4), if x is generable, then it is deducible. But, by Broad (1), if x is emergent, it is not deducible. But it then follows that if x is generable, then it cannot be emergent![31] In particular, if x is generated in an agent-based model, it cannot be classically emergent. *Agent-based modeling and classical emergentism are incompatible.* Further incompatibilities between agent-based modeling and classical emergentism will be taken up below.

Logical Confusion

Now, the logical confusion noted earlier is set forth clearly in Hempel and Oppenheim (1948), is discussed at length in Nagel (1961), and is recounted more recently by Hendriks-Jansen (1996). To summarize, like Broad, emergentists typically assert things like, "One cannot deduce higher properties from lower ones; macro properties from micro ones; the properties of the whole from the parts' properties." But, we do not deduce properties. We deduce propositions in formal languages from other propositions in those languages.[32] This is not hair-splitting: If the macro theory contains terms (predicates, variable names) that are not terms of the micro theory, then of course it is impossible to deduce macro claims involving those terms from propositions of the micro theory. It is logically impossible. So the "higher emergent"

property of water, "translucence," is trivially not deducible from the micro theory of oxygen (O) and hydrogen (H) since "translucent" is not a term of the micro theory. Many so-called "emergent properties" of "wholes" are not deducible from "parts" for this purely logical reason. So emergence, as nondeducibility, is always *relative* to some theory (some set of well-formed formulae and inference rules); it is not absolute as the classicals would have it.

A *relative* version of emergence due to Hempel and Oppenheim (1948) is formalized in Stephan (1992) as follows. Consider a system with constituents C1, . . . , Cn in relation O to one another (analogous to Broad's A, B, C, and R). "This combination is termed a *microstructure* [C1, . . . Cn;O]. And let T be a theory. Then, a system property P is emergent, *relative to* this microstructure and theory T, if:

(a.) There is a law LP which holds: for all x, when x has microstructure [C1, . . . Cn;O] then x has property P, and
(b.) By means of theory T, LP cannot be deduced from laws governing the C1, . . . Cn in isolation or in other microstructures than the given."

Stephan continues, "By this formulation the original absolute claim has been changed into a merely relative one which just states that at a certain time according to the available scientific theories we are not able to deduce the so-called emergent laws (p.39)."[33] But now, as Hempel and Oppenheim write, "If the assertion that life and mind have an emergent status is interpreted in *this* sense, then its import can be summarized approximately by the statement that no explanation, in terms of microstructure theories, is available *at present* for large classes of phenomena studied in biology and psychology." (emphases added). This quite unglamorous point, they continue, would "appear to represent the rational core of the doctrine of emergence." Not only does this relative formulation strip the term of all higher Gestalt harmonics, but it suggests that, for any given phenomenon, emergent status itself may be fleeting.

Scientific Progress

As scientific theories progress, in other words, that which was unexplainable and "emergent" ceases to be. The chemical bond—a favorite of the British emergentists—is an excellent example. Here, Terrence Horgan (1993) is worth quoting at length:

When Broad wrote, "Nothing that we know about Oxygen by itself or in its combination with anything but Hydrogen would give us the least reason to suppose that it would combine with Hydrogen at all. Nothing that we know about Hydrogen by itself or in its combinations with anything but Oxygen would give us the least reason to expect that it would combine with Oxygen at all" (1925, pp. 62–63), his claim was true. Classical physics could not explain chemical

bonding. But the claim didn't stay true for long: by the end of the decade quantum mechanics had come into being, and quantum-mechanical explanations of chemical bonding were in sight.

The chemical bond no longer seemed mysterious and "emergent." Another example was biology, for the classical emergentists a rich source of higher "emergent novelties," putatively unexplainable in physical terms. Horgan continues,

Within another two decades, James Watson and Francis Crick, drawing upon the work of Linus Pauling and others on chemical bonding. explained the information-coding and self-replicating properties of the DNA molecule, thereby ushering in physical explanations of biological phenomena in general.

As he writes, "These kinds of advances in science itself, rather than any internal conceptual difficulties, were what led to the downfall of British emergentism, as McLaughlin (1992) persuasively argues." Or, as Herbert Simon (1996) writes, "Applied to living systems the strong claim [quoting the "holist" philosopher J.C. Smuts] that 'the putting together of their parts will not produce them or account for their characters and behaviors' implies a vitalism that is wholly antithetical to modern molecular biology."

In its strong classical usage, the term "emergent" simply "baptizes our ignorance," to use Nagel's phrase (Nagel, 1961, p. 371). And, when de-mystified, it can mean nothing more than "not presently explained." But, this is profoundly different from "not explainable in principle," as Alexander and his emergentist colleagues would have it, which *is* stifling, not to mention baseless empirically. As Hempel and Oppenheim wrote,

Emergence is not an ontological trait inherent in some phenomena; rather it is indicative of the scope of our knowledge at a given time; thus it has no absolute, but a relative character; and what is emergent with respect to the theories available today may lose its emergent status tomorrow. (1948, p. 263)

Good Questions

Now, all the questions posed by agent-based modelers and complexity scientists in this connection are fine: How do individuals combine to form firms, or cities, or institutions, or ant colonies, or computing devices? These are all excellent questions. The point is that they are posable—indeed most productively posed—without the imprecise and possibly self-mystifying terminology of "emergence," or "supervenience," as Morgan called it. Obviously, "wholes" may have attributes or capabilities that their constituent parts cannot have (e.g., "whole" conscious people can have happy memories of childhood while, presumably, indi-

vidual neurons cannot). Equally obvious, the parts have to be hooked up right—or interact in specific, and perhaps complicated, ways—for the whole to exhibit those attributes.[34] We *at present* may be able to explain why these specific relationships among parts eventuate in the stated attributes of wholes, and we may not. But, unlike classical emergentists, we do not *preclude* such explanation in principle.

Indeed, by attempting to generate these very phenomena on computers or in mathematical models, we are *denying* that they are unexplainable or undeducible in principle—we're trying to explain them precisely by figuring out microrules that will generate them. In short, we agent-based modelers and complexity researchers actually part company from those, like Alexander and company, whose terminology we have, perhaps unwittingly, adopted. Lax definitions can compound the problem.

Operational Definitions

Typical of classical emergentism would be the claim: *No description of the individual bee can ever explain the emergent phenomenon of the hive.* How would one know that? Is this a falsifiable empirical claim, or something that seems true because of a lax definition of terms? Perhaps the latter. The mischievous piece of the formulation is the phrase "description of the individual bee." What is that? Does "the bee's" description not include its rules for interacting with other bees? Certainly, it makes little sense to speak of a Joshua Epstein devoid of all relationships with family, friends, colleagues, and so forth. "Man is a social animal," quoth Aristotle. My "rules of social interaction" are, in part, what make me me. And, likewise, the bee's interaction rules are what make it a bee—and not a lump. When (as a designer of agent objects) you get these rules right—when you get "the individual bee" right—you get the hive, too. Indeed, from an operationist (Hempel, 1956) viewpoint, "the bee" might be defined as that x that, when put together with other x's, *makes* the hive (the "emergent entity"). Unless the theoretical (model) bees generate the hive when you put a bunch of them together, you haven't described "the bee" adequately. Thus, contrary to the opening emergentist claim, *it is precisely the adequate description of "the individual bee" that explains the hive.* An admirable modeling effort along precisely such lines is Theraulaz *et al.* (1998).

Agent-Based Modeling Is Reductionist

Classical emergentism holds that the parts (the microspecification) cannot explain the whole (the macrostructure), while to the agent-based modeler, it is precisely the generative sufficiency of the parts (the microspecification) that constitutes the whole's explanation! In this particular sense, agent-based modeling is reductionist.[35] Classical emergentism seeks to preserve a "mystery gap" between micro and macro; agent-based modeling seeks to demystify this alleged gap by identifying microspecifications that are sufficient to generate—robustly and replicably—the macro

(whole). Perhaps the following thoughts of C.S. Peirce (1879) are apposite:

> One singular deception . . . which often occurs, is to mistake the sensation produced by our own unclearness of thought for a character of the object we are thinking. Instead of perceiving that the obscurity is purely subjective, we fancy that we contemplate a quality of the object which is essentially mysterious; and if our conception be afterward presented to us in a clear form we do not recognize it as the same, owing to the absence of the feeling of unintelligibility.

Explanation and Prediction

A final point is that classical emergentism traffics on a crucial (and to this day quite common) confusion: between explanation and prediction. It may well be that certain phenomena are unpredictable in principle (e.g., stochastic). But that does not mean—as classical emergentists would have it—that they are unexplainable in principle. Plate tectonics explains earthquakes but does not predict their occurrence; electrostatics explains lightning but does not predict where it will hit; evolutionary theory explains species diversity but does not predict observed phenotypes. In short, one may grant unpredictability without embracing "emergence," as absolute unexplainability, à la Alexander and colleagues.[36] And, of course, it may be that in some cases prediction is a perfectly reasonable goal. (For further distinctions between prediction and explanation, see Scheffler, 1960; Suppes, 1985; and Newton-Smith, 1981).

In its strong classical usages—connoting absolute nondeducibility and absolute unexplainability—"emergentism" is logically confused and antiscientific. In weak level-headed usages—like "arising from local agent interactions"—a special term hardly seems necessary. For other attempts to grapple with the term "emergent," see Cariani (1992), Baas (1994), Gilbert (1995), and Darley (1996). At the very least, practitioners—and I include myself—should define this term carefully when they use it and distinguish their, perhaps quite sensible, meaning from others with which the term is strongly associated historically. To anyone literate in the philosophy of science, "emergence" has a history, and it is one with that many scientists may—indeed should—wish to part company. Doubtless, my own usage has been far too lax, so this admonition is directed as much at myself as at colleagues.

14. RECAPITULATION AND CONCLUSION

I am not a soldier in an agent-based methodological crusade. For some explanatory purposes, low dimensional differential equations are perfect. For others, aggregate regression is appropriate. Game theory offers deep insight in nu-

merous contexts and so forth. But agent-based modeling is clearly a powerful tool in the analysis of *spatially distributed systems of heterogeneous autonomous actors with bounded information and computing capacity.* It is the main scientific instrument in a generative approach to social science, and a powerful tool in empirical research. It is well suited to the study of connectionist phenomena in social science. It offers a natural environment for the study of certain interdisciplinary questions. It allows us to test the sensitivity of theories, such as neoclassical microeconomics, to relaxations in core assumptions (e.g., the assumption of fixed preferences). It allows us to trace how individual (micro) rules generate macroscopic regularities. In turn, we can employ laboratory behavioral research to select among competing multiagent models having equal generative power. The agent-based approach may decouple individual rationality from macroscopic equilibrium and separate decision science from social science more generally. It invites a synthesis of analytical and computational perspectives that is particularly relevant to the study of non-equilibrium systems. Agent-based models have significant pedagogical value. Finally, the computational interpretation of social dynamics raises foundational issues in social science—some related to intractability, and some to undecidability proper. Despite a number of significant challenges, agent-based computational modeling can make major contributions to the social sciences.

ACKNOWLEDGMENTS

NOTES

1. A basic exposure to agent-based computational modeling—or artificial societies—is assumed. For an introduction to agent-based modeling and a discussion of its intellectual lineage, see Epstein and Axtell (1996). I use the term "computational" to distinguish artificial societies from various equation-based models in mathematical economics, n-person game theory, and mathematical ecology that (while not computational) can legitimately be called agent-based. These equation-based models typically lack one or more of the characteristic features of computational agent models noted below. Equation-based models are often called "analytical" (as distinct from computational), which occasions no confusion so long as one understands that "analytical" does not mean analytically tractable. Indeed, computer simulation is often needed to approximate the behavior of particular solutions. The relationship of agent-based models and equations is discussed further below.

2. The term "emergence" and its history are discussed at length below. Here, I use the term "emergent" as defined in Epstein and Axtell (1996, p. 35), to mean simply "arising from the local interaction of agents."

3. The features noted here are not meant as a rigid definition; not all agent-based models exhibit all these features. Hence, I note that the exposition is in a strong form. The point is that these characteristics are easily arranged in agent-based models.

4. See Kirman (1992).

5. For analytical models of local interactions, see Blume and Durlauf (1998).

6. See Simon (1982) and Rubinstein (1998).

7. We will refer to an initial agent-environment specification as a microspecification. While, subject to outright computational constraints, agent-based modeling *permits* extreme methodological individualism, the "agents" in agent-based computational models are not always individual humans. Thus, the term "microspecification" implies substantial—but not necessarily complete—disaggregation. Agent-based models are naturally implemented in object-oriented programming languages in which agents and environmental sites are objects with fixed and variable internal states (called instance variables), such as location or wealth, and behavioral rules (called methods) governing, for example, movement, trade, or reproduction. For more on software engineering aspects of agent-based modeling, see Epstein and Axtell (1996).

8. Though he does not use this terminology, Schelling's (1971) segregation model is a pioneering example.

9. In fact, Schelling did his early experiments without a computer. More to the point, one might argue that, for example, Uzawa's (1962) analytical model of non-equilibrium trade in a population of agents with heterogeneous endowments is generative.

10. Likewise, it would be wrong to claim that Arrow-Debreu general equilibrium theory is devoid of explanatory power because it is not generative. It addresses different questions than those of primary concern here.

11. For expository purposes, I write as though a macrostructure is either generated or not. In practice, it will generally be a question of degree.

12. Locating this (admittedly informal) usage of "explanation" in the vast and contentious literature on that topic is not simple and requires a separate essay. For a good collection on scientific explanation, see Pitt (1988). See also Salmon (1984), Cartwright (1983), and Hausman (1992). Very briefly, because no general scientific (covering) laws are involved, generative sufficiency would clearly fail one of Hempel and Oppenheim's (1948) classic deductive-nomological requirements. Perhaps surprisingly, however, it meets the deduction requirement itself, as shown by the Theorem below. That being the case, the approach would appear to fall within the hypothetico-deductive framework described in Hausman (1992, p. 304). A microspecification's failure to generate a macrostructure falsifies the hypothesis of its sufficiency and disqualifies it as an explanatory candidate, consistent with Popper (1959). Of course, sorting out exactly what component of the microspecification—core agent rules or auxiliary conditions—is producing the generative failure is the Duhem problem. Our weak requirements for explanatory candidacy would seem to have much in common with the *constructive empiricism* of van Fraassen (1980). On this antirealist position, truth (assuming it has been acceptably defined) is eschewed as a goal. Rather, "science aims to give us theories which are empirically adequate; and acceptance of a theory involves as belief only that it is empirically adequate." (von Fraassen, 1980, p. 12.) However, faced with competing microspecifications that are *equally adequate* empirically (i.e., do equally well in generating a macro target), one would choose by the criterion of empirical plausibility at the micro level, as determined experimentally. On realism in social science, see Hausman (1998).

13. Constructivism in this mathematical sense should not be confused with the doctrine of social constructionism sometimes identified with so-called "post-modernism" in other fields.

14. In such applications, it may be accurate to speak of an inductive statistical (see Salmon, 1984) account over many realizations, each one of which is, technically, a deduction (by the Theorem above).

15. An agent-based model can be interpreted as furnishing a kind of *constructive* existence proof. See Axelrod (1997).

16. See Chomsky (1965, p.3). The "syntactic component of a generative grammar," he writes, is concerned with "rules that specify the well-formed strings of minimal syntactically functioning units . . . " I thank Samuel David Epstein for many fruitful discussions of this parallel.

17. More precisely, for each candidate rule (or agent specification), one runs a large population of simulated histories—each with its own random seed. The question then becomes: where, in the population of simulated histories is the true history? Rules that generate distributions with the true history (i.e., it's statistic) at the mean enjoy more explanatory power than rules generating distributions with the true history at a tail.

18. I thank Robert Axtell for this term.

19. In a different context, the sensitivity to network topology is studied computationally by Bagley and Farmer (1992).

20. This is shorthand for the appropriate procedure in which one would generate distributions of outcomes for the two assumptions and test the hypothesis that these are indistinguishable statistically.

21. For deep work on gene-culture co-evolution generally, using different techniques, see Feldman and Laland (1996).

22. More precisely, micro rationality is not necessary for *some* equilibrium, but it may a different equilibrium from the one that would occur were agents rational.

23. Again, as a policy application, agent-based modeling might suggest ways to operate on—or "tip"—ethnic animosity itself.

24. See, for example, Hirsch, et al. (1989).

25. The important Arrow Impossibility Theorem (Arrow, 1963) strikes me as different in nature from these sorts of results. It does not turn—as these results do—on the existence of sets that are recursively enumerable but not recursive.

26. Here, we are discussing regularities in model output alone, not the relationship of model output to some real-world data set, as in the Anasazi project.

27. As we wrote there, "A particularly loose usage of 'emergent' simply equates it with 'surprising,' or 'unexpected,' as when researchers are unprepared for the kind of systematic behavior that emanates from their computers." We continued, "This usage obviously begs the question, 'Surprising to whom?'." (Epstein and Axtell, 1996, p. 35.)

28. Although many contemporary researchers do not use the term in this way, others assume that this is the generally accepted meaning. For example, Jennings, Sycara, and Woolridge (1998) write that " . . . the very term 'emerges' suggests that the relationship between individual behaviors, environment, and overall behavior is not understandable," which is entirely consistent with the classical usage.

29. To highlight the chasm between this classical and certain modern usages, while Broad defines emergence as undeducible, Axelrod (1997c, p. 194) writes that "there are some models . . in which emergent properties can be formally deduced."

30. For example, note that we can deduce Alexander's Law (2) from the others. By Broad (1), if x is classically emergent then it is not deducible; but then by (4) and *modus tollens,* x is not generable; and then by the Motto (3), x is not explainable. So, by hypothetical syllogism, we obtain Alexander (2). In a punctilious derivation, we would of course invoke universal intantiation first; then rules of an explicit sentential calculus (e.g., Copi, 1979), and then use universal generalization.

31. Filling in for any x, (4) is Gx ⊃ Dx; but Dx = ¬(¬Dx), from which ¬Cx follows from (1) by *modus tollens.*

32. Formal systems are closed under their rules of inference (e.g., *modus ponens*) in the sense that propositions in a formal system can only be deduced from other propositions of that system.

33. Contemporary efforts (see Baas, 1994) to define a kind of relative (or hierarchical) emergence by way of Godel's First Theorem (see Smullyan, 1992) seem problematic. Relative to a given (consistent and finitely axiomatized) theory, T, Baas calls undecidable sentences "observationally emergent." However, we are presumably interested in generating "emergent phenomena" in computational models. And it is quite unclear from what *computational* process Baas's observationally emergent entities—undecidable sentences of T—would actually emerge since, by Tarski's Theorem, the set of true and undecidable propositions is not recursively enumerable. (Hodel, 1995, p. 310, p. 354).

34. There is no reason to present these points as if they were notable, as in the following representative example: " . . . put the parts of an aeroplane together in the correct relationship and you get the emergent property of flying, even though none of the parts can fly" Johnson (1995; p. 26).

35. The term "reductionist" admits a number of definitions. We are not speaking here of the reduction of *theories,* as in the reduction of thermodynamics to statistical mechanics. See Nagel (1961), Garfinkel (1991), and Anderson (1972).

36. On fundamental sources of unpredictability, see Gell-Mann (1997).

REFERENCES

1. Akerlof, G. 1997. Social Distance and Social Decisions. Econometrica 65: 1005–1027.

2. Albin, P.; Foley, D.K. 1990. Decentralized, Dispersed Exchange without an Auctioneer: A Simulation Study. Journal of Economic Behavior and Organization 18(1).

3. Albin, P.S.; Foley, D.K. 1997. The Evolution of Cooperation in the Local-Interaction Multi-Person Prisoners' Dilemma. Working Paper, Department of Economics, Barnard College, Columbia University.

4. Albin, P.S. 1998. Barriers and Bounds to Rationality. Princeton: Princeton University Press.

5. Alexander, S. 1920. Space, Time, and Deity: The Gifford Lectures at Glasgow, 1916–1918. New York: Dover.

6. Anderson, P.W. 1972. More Is Different. Science 177: 393–396.

7. Arrow, K. 1963. Social Choice and Individuxal Values. New York: Wiley.

8. Arthur, W.B.; LeBaron, B.; Palmer, R.; Tayler, P. 1997. Asset Pricing under Endogenous Expectations in an Artificial Stock Market. In The Economy as a Complex Evolving System II. Arthur, W.B.; Durlauf, S.; Lane, D. eds. Menlo Park: Addison-Wesley.

9. Ashlock, D.; Smucker, D.M.; Stanley, E.A.; Tesfatsion, L. 1996. Preferential Partner Selection in an Evolutionary Study of Prisoner's Dilemma. BioSystems 37: 99–125.

10. Axelrod, R. 1987. The Evolution of Strategies in the Iterated Prisoner's Dilemma. In Davis, L. ed. Genetic Algorithms and Simulated Annealing. London: Pitman: 32–41.

11. Axelrod, R.; Bennett, D.S. 1993. A Landscape Theory of Aggregation. British Journal of Political Science 23:211–233.

12. Axelrod, R. 1995. A Model of the Emergence of New Political Actors. In Gilbert and Conte, eds. Reprinted in Axelrod 1997.

13. Axelrod, R. 1997. The Dissemination of Culture: A Model with Local Convergence and Global Polarization. Journal of Conflict Resolution 41: 203–26. Reprinted in Axelrod 1997b.

14. Axelrod, R. 1997b. The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton: Princeton University Press.

15. Axelrod, R. 1997c. Advancing the Art of Simulation in the Social Sciences, Complexity 3: 193–199.

16. Axtell, R. 1999. The Emergence of Firms in a Population of Agents: Local Increasing Returns, Unstable Nash Equilibria, and Power Law Size Distributions. Santa Fe Institute Working Paper 99-03-019.

17. Axtell, R.L.; Epstein, J.M. 1994. Agent-Based Modeling: Understanding Our Creations. Bulletin of the Santa Fe Institute 9: 28–32.

18. Axtell, R.; Epstein, J.M. 1999. Coordination in Transient Social Networks: An Agent-Based Computational Model of the Timing of Retirement. In Aaron, H., ed. Behavioral Dimensions of Retirement Economics. Russell Sage.

19. Axtell, R.; Axelrod, R.; Epstein, J.M.; Cohen, M.D. 1996. Aligning Simulation Models: A Case Study and Results. Computational and Mathematical Organization Theory 1:123–141.

20. Axtell, R.; Epstein, J.M.; Young, H.P. 1998. The Emergence of Economic Classes in an Agent-Based Bargaining Model. Forthcoming in Durlauf, S.; Young, H.P. eds. Social Dynamics. Cambridge: MIT Press.

21. Baas, N.A. 1994. Emergence, Hierarchies, and Hyperstructures. In Langton, C.G., ed. Artificial Life III. Reading; Addison-Wesley. pp. 515–537.

22. Bagley, R.J.; Farmer, J.D. 1992. Spontaneous Emergence of a Metabolism. In Langton, *et al.* 1992.

23. Bak, P.; Chen, K.; Scheinkman, J.; Woodford, M. 1993. Aggregate Fluctuations from Independent Sectoral Shocks: Self-Organized Criticality in a Model of Production and Inventory Dynamics. Ricerche Economiche 47: 3–30.

24. Bak, P.; Paczuski, M.; Shubik, M. 1996. Price Variations in a Stock Market with Many Agents. Santa Fe Institute Working Paper 96-09-075.

25. Barnsley, M. 1988. Fractals Everywhere. San Diego: Academic Press.

26. Ben-Porath, E. 1990. The Complexity of Computing Best Response Automata in Repeated Games with Mixed Strategies. Games and Economic Behavior 2: 1–12.

27. Bicchieri, C.; Jeffrey, R.; Skyrms, B. eds. 1997. The Dynamics of Norms. New York: Camridge University Press.

28. Binmore, K.G. 1987. Modeling Rational Players, I. Economics and Philosophy 3: 179–214.

29. Binmore, K.G. 1988. Modeling Rational Players, II. Economics and Philosophy 4: 9–55.

30. Blume, L.E.; Durlauf, S. 1998. The Interactions-Based Approach to Socioeconomic Behavior. Forthcoming in Durlauf S.; Young, H.P. eds. Social Dynamics. Cambridge: MIT Press.

31. Bowles, S. 1998. Endogenous Preferences: The Cultural Consecuences of Markets and Other Institutions. Journal of Economic Literature XXXVI: 75–111.

32. Broad, C.D. 1925. The Mind and its Place in Nature. London: Routledge and Kegan Paul, Ltd.

33. Bullard, J.; Duffy, J. 1998. Learning and Excess Volatility. Macroeconomics Seminar Paper, Federal Reserve Bank of St. Louis.

34. Buss, S.; Papadimitriou, C.H.; Tsitsiklis, J.N. 1991. On the Predictability of Coupled Automata. Complex Systems 2: 525–539.

35. Camerer, C.F. 1997. Progress in Behavioral Game Theory. Journal of Economic Perspectives. 11: 167–188.

36. Camerer, C.F.; Thaler, R. 1995. Anomalies: Ultimatums, Dictators, and Manners. Journal of Economic Perspectives 9: 2, 209–219.

37. Cariani, P. 1992. Emergence and Artificial Life. In Langton, *et al.* 1992.

38. Cartwright, N. 1983. How the Laws of Physics Lie. New York: Oxford University Press.

39. Cederman, L.-E. 1997. Emergent Actors in World Politics: How States and Nations Develop and Dissolve. Princeton: Princeton University Press.

40. Chalmers, A.F. 1982. What Is This Thing Called Science? Cambridge: Hackett.

41. Chomsky, N. 1965. Aspects of the Theory of Syntax. Cambridge, MA: MIT Press.

42. Copi, I. 1979. Symbolic Logic. New York: Macmillan.

43. Crutchfield, J.P.; Mitchell, M. 1995. The Evolution of Emergent Computation. Proceedings of The National Academy of Sciences. 92: 1074–10746.

44. Darley, O. 1996. Emergent Phenomena and Complexity. Artificial Life V. Cambridge, MA: MIT Press.

45. Dean, J.S.; Gumerman, G.J.; Epstein, J.M.; Axtell, R.; Swedlund, A.C.; Parker, M.T.; McCarroll, S. 1999. Understanding Anasazi Culture Change Through Agent-Based Modeling. Forthcoming in Gumerman, George J.; Kohler T., eds. Modeling Small Scale Societies. New York: Oxford University Press.

46. Durlauf, S.N. 1997. Statistical Mechanics Approaches to Socioeconomic Behavior. In The Economy as a Complex Evolving System II, Arthur, W.B.; Durlauf, S.; Lane, D., eds. Menlo Park: Addison-Wesley.

47. Durlauf, S.N. 1997. Limits to Science or Limits to Epistemology? Complexity 2: 31–37.

48. Epstein, J.M. 1997. Nonlinear Dynamics, Mathematical Biology, and Social Science. Menlo Park: Addison-Wesley.

49. Epstein, J.M. 1998. Zones of Cooperation in Demographic Prisoner's Dilemma. Complexity 4: 2, 36–48.

50. Epstein, J.M.; Axtell, R. 1996. Growing Artificial Societies: Social Science From the Bottom Up. Cambridge, MA: MIT Press.

51. Epstein, J.M.; Axtell, R. 1997. Artificial Societies and Generative Social Science. Artificial Life and Robotics 1:33–34.

52. Feldman, M.W.; Laland, K.N. 1996. Gene-Culture Coevolutionary Theory. Trends in Ecology and Evolution 11: 453–457.

53. Feldman, M.W.; Cavalli-Sforza, L.L. 1977. The Evolution of Continuous Variation II: Complex Transmission and Assortative Mating. Theoretical Population Biology 11: 161–181.

54. Feldman, M.W.; Cavalli-Sforza, L.L. 1976. Cultural and Biological Evolutionary Processes: Selection for a Trait under Complex Transmission. Theoretical Population Biology 9: 239–259.

55. Flierl, G.; Grünbaum, D.; Levin, S.A.; Olson, D. 1999. From Individuals To Aggregations: The Interplay Between Behavior and Physics. J. Theoretical Biology. 196: 397–454.

56. Fraenkel, A.A.; Bar-Hillel, Y. 1958. Foundations of Set Theory. Amsterdam: North-Holland.

57. Friedman, M. 1953. The Methodology of Positive Economics. In Essays in Positive Economics. Chicago: University of Chicago Press.

58. Garfinkel, A. Reductionism. In Boyd, R.; Gasper, P.; Trout, J.D. eds. 1991. The Philosophy of Science. Cambridge, MA: MIT Press.

59. Gell-Mann, M. 1997. Fundamental Sources of Unpredictability. Complexity 3: 1, 9–13.

60. Gilbert, N. 1995. Emergence in Social Simulation. In Gilbert, N.; Conte, R. 144–156.

61. Gilbert, N.; Conte, R. eds. 1995. Artificial Societies: The Computer Simulation of Social Life. London: University College Press.

62. Gilboa, I. 1988. The Complexity of Computing Best-Response Automata in Repeated Games. Journal of Economic Theory 45: 342–352.

63. Glaeser, E.; Sacerdote, B.; Scheinkman, J. 1996. Crime and Social Interactions. Quarterly Journal of Economics 111: 507–548.

64. Glymour, C. 1980. Theory and Evidence. Princeton: Princeton University Press.

65. Hausman, D.M. 1992. The Inexact and Separate Science of Economics. Cambridge: Cambridge University Press.

66. Hausman, D.M. 1998. Problems with Realism in Economics. Economics and Philosophy 14: 2.

67. Hempel, C.G.; Oppenheim, P. 1948. Studies in the Logic of Explanation. Philosophy of Science 15: 567–579.

68. Hempel C.G. 1956. A Logical Appraisal of Operationism. In Frank, P., ed, The Validation of Scientific Theories. Boston: Beacon Press.

69. Hendriks-Jansen, H. 1996. In Praise of Interactive Emergence, Or Why Explanations Don't Have to Wait for Implementation. In Boden, M.A., ed., The Philosophy of Artificial Life. New York: Oxford University Press.

70. Hirsch, M.D.; Papadimitriou, C.H.; Vavasis, S.A. 1989. Exponential Lower Bounds for Finding Brouwer Fixed Points. Journal of Complexity 5: 379–416.

71. Hodel, R.E. 1995. An Introduction to Mathematical Logic. Boston: PWS Publishing Company.

72. Holland, J.H. 1992. Genetic Algorithms. Scientific American 267: 66–72.

73. Horgan, T. 1993. From Supervenience to Superdupervenience: Meeting the Demands of a Material World. Mind 102: 555–586.

74. Howson, C.; Urbach, P. 1993. Scientific Reasoning: A Bayesian Approach. Chicago: Open Court.

75. Huberman, B.; Glance, N. 1993. Evolutionary Games and Computer Simulations. Proceedings of the National Academy of Sciences 90: 7715–7718.

76. Ijiri, Y.; Simon, H. 1977. Skew Distributions and the Sizes of Business Firms. New York: North-Holland.

77. Ilachinski, A. 1997. Irreducible Semi-Autonomous Adaptive Combat (ISAAC): An Artificial-Life Approach to Land Warfare. Center for Naval Analyses Research Memorandum CRM 97-61.10.

78. Jeffrey, R. 1991 (3rd edition). Formal Logic: Its Scope and Limits. New York: McGraw-Hill.

79. Jennings, N.; Sycara, K; Woolridge, M. 1998. A Roadmap of Agent Research and Development. Autonomous Agents and Multi-Agent Systems: 7–38.

80. Johnson, J. 1995. A Language of Structure in the Science of Complexity. Complexity 1: 3, 22–29.

81. Kagel, J.H.; Roth, A.E. eds. 1995. The Handbook of Experimental Economics. Princeton: Princeton University Press.

82. Kalai, E. 1990. Bounded Rationality and Strategic Complexity in Repeated Games. In Ichiishi, T.; Neyman, A.; Tauman Y. eds. Game Theory and Applications. San Diego: Academic Press.

83. Katona, G. 1951. Psychological Analysis of Economic Behavior. New York: McGraw-Hill.

84. Kauffman, S.A. 1993. The Origins of Order: Self-Organization and Selection in Evolution. New York: Oxford University Press.

85. Kirman, A. 1992. Whom or What Does the Representative Individual Represent. Journal of Economic Perspectives 6: 117–136.

86. Kirman, A.P.; Vriend, N.J. 1998. Evolving Market Structure: A Model of Price Dispersion and Loyalty. Working Paper. University of London, London, UK.

87. Kleene, S.C. 1967. Mathematical Logic. New York; Wiley.

88. Knuth, D.E. 1969. The Art of Computer Programming, Volume 2: Seminumerical Algorithms. Reading: Addison-Wesley.

89. Kollman, K.; Miller, J.; Page S. 1992. Adaptive Parties in Spatial Elections. American Political Science Review 86: 929–937.

90. Kuran, T. 1989. Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution. Public Choice 61: 41–74.

91. Langton, C.G.; Taylor, C.; Farmer, J.D.; Rasmussen, S. 1992. Artificial Life II. New York: Addison-Wesley.

92. Lewis, A.A. 1985. On Effectively Computable Realizations of Choice Functions. Mathematical Social Sciences 10: 43–80.

93. Lewis, A.A. 1992a. Some Aspects of Effectively Constructive Mathematics that Are Relevant to the Foundations of Neoclassical Mathematical Economics and the Theory of Games. Mathematical Social Sciences 24: 209–235.

94. Lewis, A.A. 1992b. On Turing Degrees of Walrasian Models and a General Impossibility Result in the Theory of Decision Making. Mathematical Social Sciences 25.

95. Lindgren, K.; Nordahl, M.G. 1994a. Cooperation and Community Structure in Artificial Ecosystems. Artificial Life 1: 15–37.

96. McLaughlin, B.P. 1992. The Rise and Fall of British Emergentism. In Beckermann, A.; Flohr, H.; Kim, J. eds. Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism. New York: Walter de Gruyter.

97. McNeill, W.H. 1976. Plagues and Peoples. Anchor Press/Doubleday.

98. Miller, J.H. 1996. The Coevolution of Automata in the Repeated Prisoner's Dilemma. Journal of Economic Behavior and Organization 29: 87–112.

99. Minsky, M. 1985. The Society of Mind. New York: Simon and Schuster.

100. Mitchell, M. 1998. An Introduction to Genetic Algorithms. Cambridge, MA: MIT Press.

101. Morgan, C.L. 1923. Emergent Evolution. London: Williams and Norgate.

102. Nagel, E. 1961. The Structure of Science. New York: Harcourt, Brace, and World.

103. Nagel, K.; Rasmussen, S. 1994. Traffic at the Edge of Chaos. In Brooks, R., ed., Artificial Life IV. Cambridge, MA: MIT Press.

104. Newton-Smith, W.H. 1981. The Rationality of Science. New York: Routledge.

105. Neyman, A. 1985. Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma. Economics Letters 19: 227–229.

106. Nordhaus, W.D. 1992. Lethal Model 2: The Limits to Growth Revisited. Brookings Papers on Economic Activity: 2: 1–59.

107. Nowak, M.A.; May, R.M. 1992. Evolutionary Games and Spatial Chaos. Nature 359: 826–829.

108. Papadimitriou, C. 1993. Computational Complexity as Bounded Rationality. Mathematics of Operations Research.

109. Papadimitriou, C.H.; Yannakakis, M. 1994. On Complexity as Bounded Rationality. Proceedings of the Twenty-Sixth Annual ACM Symposium on the Theory of Computing, pp. 726–733. Montreal, Quebec, Canada, 23–25 May 1994.

110. Peirce, C.S. 1879. How to Make Our Ideas Clear. In Epstein, J.; Kennedy G. eds., 1967, The Process of Philosophy: A Historical Introduction. New York: Random House. pp. 770–790.

111. Pitt, J.C. ed. 1988. Theories of Explanation. New York: Oxford University Press.

112. Popper, K.R. 1959. The Logic of Scientific Discovery. New York: Routledge.

113. Popper, K.R. 1963. Conjectures and Refutations: The Growth of Scientific Knowledge. New York: Routledge.

114. Prasad, K. 1997. On the Computability of Nash Equilibria. Journal of Economic Dynamics and Control 21: 943–953.

115. Prietula, M.J.; Carley, K.M.; Gasser, L. 1998. Simulating Organizations. Cambridge, MA: MIT Press.

116. Rabin, M. 1998. Psychology and Economics. Journal of Economic Literature 36: 11–46.

117. Rabin, M.O. 1957. Effective Computability of Winning Strategies. In Dresher, M.; Tucker, A.W.; Wolfe, P. eds. Contributions to the Theory of Games, Annals of Mathematical Studies 3 (39). Princeton: Princeton University Press.

118. Resnick, M. 1994. Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds. Cambridge, MA: MIT Press.

119. Rogers, H. 1967. Theory of Recursive Functions and Effective Computability. New York: McGraw-Hill.

120. Rubinstein, A. 1986. Finite Automata Play the Repeated Prisoners' Dilemma. Journal of Economic Theory 39: 83–96.

121. Rubinstein, A. 1998. Modeling Bounded Rationality. Cambridge, MA: MIT Press.

122. Salmon, W.C. 1984. Scientific Explanation and the Causal Structure of the World. Princeton: Princeton University Press.

123. Salmon, W.C. with Jeffrey, R.C.; Greeno, J.G. 1971. Statistical Explanation and Statistical Relevance. Pittsburgh: University of Pittsburgh Press.
124. Scheffler, I. 1960. Explanation, Prediction, and Abstraction. In Danto, A.; Morgenbesser, S. Philosophy of Science. New York: Meridian Books.
125. Schelling, T. 1971. Dynamic Models of Segregation. Journal of Mathematical Sociology 1: 143–186.
126. Schelling, T.C. 1978. Micromotives and Macrobehavior. New York: W.W. Norton and Company.
127. Simon, H.A. 1982. Models of Bounded Rationality. Cambridge, MA: MIT Press.
128. Simon, H. 1978. On How to Decide What to Do. The Bell Journal of Economics.
129. Simon, H. 1996 (3rd edition). The Sciences of the Artificial. Cambridge, MA: MIT Press.
130. Skyrms, B. 1986, 3rd ed. Choice and Chance, Belmont, CA: Wadsworth.
131. Skyrms, B. 1998. Evolution of the Social Contract. Cambridge: Cambridge University Press.
132. Smullyan, R.M. 1992. Godel's Incompleteness Theorems. New York: Oxford University Press.
133. Sober, E. 1996. Learning from Functionalism—Prospects for Strong Artificial Life. In Boden, M.A. ed., The Philosophy of Artiicial Life. New York: Oxford University Press.
134. Stanley, M.H.R.; Amaral, L.A.N.; Buldyrev, S.V.; Havlin, S.; Leschhorn, H.; Maass, P.; Salinger, M.A.; Stanley, H.E. 1996. Scaling Behavior in the Growth of Companies. Nature, 379: 804–806.
135. Stephan, A. 1992. Emergence—A Systematic View on its Historical Facets. In Beckermann, A.; Flohr, H.; Kim, J. eds. Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism. New York: Walter de Gruyter.
136. Suppes, P. 1985. Explaining the Unpredictable. Journal of Philosophy.
137. Tesfatsion, L. 1995. A Trade Network Game with Endogenous Partner Selection. Economic Report 36. Iowa State University, Department of Economics.
138. Theraulaz, G.; Bonabeau, E.; Deneubourg, J.-L. The Origin of Nest Complexity in Social Insects. Complexity 3: 6, 15–25.
139. Topa, G. 1997. Social Interactions, Local Spillovers and Unemployment. Manuscript, Department of Economics, New York University.
140. Tversky, A; Kahneman, D. 1986. Rational Choice and The Framing of Decisions. In Hogarth, R.M.; Reder, M.W. eds. Rational Choice: The Contrast between Economics and Psychology. Chicago: University of Chicago Press.
141. Uzawa, H. 1962. On the Stability of Edgeworth's Barter Process. International Economic Review 3: 218–232.
142. van Fraassen, B.C. 1980. The Scientific Image. New York: Oxford University Press.
143. Watkins, J.W.N. 1957. Historical Explanation in the Social Sciences. British Journal for the Philosophy of Science 8: 106.
144. Young, H.P. 1995. The Economics of Convention. Journal of Economic Perspectives 10: 105–122.
145. Young, H.P. 1993. The Evolution of Conventions. Econometrica 61: 57–84.
146. Young, H.P.; Foster, D. 1991. Cooperation in the Short and in the Long Run. Games and Economic Behavior 3: 145–156.
147. Young, H.P. 1998. Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton: Princeton University Press.