

THE POTENTIAL AND LIMITATIONS OF AGENT-BASED SIMULATION

An introduction

Gianluca Manzo

Translated by Toby Matthews

Presses de Sciences Po | « *Revue française de sociologie* »

2014/4 Vol. 55 | pages 653 - 688

ISSN 0035-2969

ISBN 9782724633771

This document is a translation of:

Gianluca Manzo and Translated by Toby Matthews, « Potentialités et limites de la simulation multi-agents : une introduction », *Revue française de sociologie* 2014/4 (Vol. 55), p. 653-688.

DOI 10.3917/rfs.554.0653

Available online at :

http://www.cairn-int.info/article-E_RFS_554_0653--the-potential-and-limitations-of.htm

The English version of this issue is published thanks to the support of the CNRS

How to cite this article :

Gianluca Manzo, « Potentialités et limites de la simulation multi-agents : une introduction », *Revue française de sociologie* 2014/4 (Vol. 55), p. 653-688.

DOI 10.3917/rfs.554.0653

Electronic distribution by Cairn on behalf of Presses de Sciences Po.

© Presses de Sciences Po. All rights reserved for all countries.

Reproducing this article (including by photocopying) is only authorized in accordance with the general terms and conditions of use for the website, or with the general terms and conditions of the license held by your institution, where applicable. Any other reproduction, in full or in part, or storage in a database, in any form and by any means whatsoever is strictly prohibited without the prior written consent of the

publisher, except where permitted under French law.

The Potential and Limitations of Agent-based Simulation

An Introduction

Gianluca MANZO

Abstract This article has a general goal: it provides an overview of the main theoretical and methodological dimensions structuring the field of research that is based on agent-based modelling and simulation. It also has a more specific aim, namely to help the reader better understand the contributions in the special issue the *Revue Française de Sociologie* is devoting to this method. To meet both requirements, the article addresses the following points: 1) the originality of agent-based modelling and simulation; 2) the kind of explanation facilitated by this method; 3) the method's computational root; 4) its flexibility in dealing with the "micro-macro transition" problem; 5) the link that can be established between agent-based modelling and empirical data; 6) the criticisms this method usually receives.

Key words. ACTION THEORY—AGENT-BASED SIMULATION—CALIBRATION—CELLULAR AUTOMATA—EXPLANATION—GENERATIVE MODELS—MATHEMATICAL MODELS—MICRO-MACRO TRANSITION—MICROSIMULATION—NUMERICAL SIMULATION—ROBUSTNESS ANALYSIS—SENSITIVITY ANALYSIS—SOCIAL INTERACTIONS—SOCIAL MECHANISMS—VALIDATION

To introduce the subject of this article I propose to begin with an example. Imagine having identified on a macro-scale a positive relationship between actors' socio-economic backgrounds and their chances of reaching a higher educational level (Jackson, 2012). Suppose also that we want to explain this relationship and that to do so we hypothesize that this association is derived from a complex set of cultural factors and strategic decisions (Jonsson and Erikson, 2000). To test this hypothesis the researcher can choose from several research strategies (Kroneberg and Kalter, 2012). For example, data could be collected on a range of cultural and economic resource indicators and, using multivariate statistics, the researcher could study whether there was a systematic variation between the distribution of these resources between socio-occupational groups and the preferences among actors for a particular level of education (Stocké, 2008). The researcher could also carry out laboratory experiments to assess the way in which actors make educational decisions when certain parameters of the decision structure, such as the benefits and risks associated with different educational options, are artificially changed (Page, Garboua and Montmarquette, 2007). Another option is to recreate the relationship between the actors' social groups and educational levels within a virtual population in which,

Translated by Toby Matthews

I would like to thank Alexandra Frénod for having completely reread the original article and refined the French, and Toby Matthews for translating it into English.

in the same way as a video game, each entity evolves over time on the basis of rules of behaviour and interaction established by the researcher (Manzo, 2013a). If one opts for this research strategy one is embarking on an agent-based simulation, which is the method I intend to present in this article.

It should be acknowledged from the outset that this task is not an easy one given that a substantial amount of work on modelling and agent-based simulation has now been done. Nearly ten years after the first literature reviews (Macy and Willer, 2002; Sawyer, 2003), it seems impossible to cover all the debates currently shaping the research domain in a single article (for a recent work that attempts this, see Edmonds and Meyer, 2013). This is the result of several developments. Firstly, there has been an increasing number of social phenomena that have been studied using agent-based simulation (for an overview, see Squazzoni, 2012). Secondly, there has been a diversification in the disciplines that use this method (in economics, see Farmer and Foley, 2009; in finance, Mathieu, Beaufls and Brandouy, 2005; in political science, Cederman, 2005; in geography, O'Sullivan, 2008; in criminology, Birks, Townsley and Stewart, 2012; in epidemiology, Auchincloss and Roux, 2008; in social psychology, Smith and Conrey, 2007; in demography, Billari and Prskawetz 2003; in biology, Thorne, Bailey and Peirce, 2007). Thirdly, agent-based simulation is now often used in combination with other more classical methods, such as statistical analysis, network analysis and laboratory experiments (for a general discussion, see Chattoe-Brown, 2014). Fourthly, these developments have, in turn, stimulated deep reflection in the philosophy of science (see, among others, Grüne-Yanoff, 2009; Marchionni and Ylikoski, 2013). Finally, agent-based simulation now plays a part in interdisciplinary research, such as socio-physics (Helbing, 2012) and studies on “complexity” (Miller and Page, 2007).

In this introductory article I intend to present the themes that run through the varied literature on agent-based simulation, focusing on the aspects that may be of particular interest to sociologists. My aim is to reveal both the potential and the limitations of the method. To this end, I will clarify what agent-based simulation is and in what ways it differs from other older forms of computer simulation currently used in our discipline. Then I will address the problem of the type of explanation made possible through agent-based simulation. In order to justify the claims made in these first two parts, I will ask the reader to follow me in an analysis of the method's low-level infrastructure. On this basis, I will discuss the classic problem in sociology of the “micro–macro” transition and explain how agent-based simulation can make a contribution. Next I will discuss the relations that can be established between agent-based simulation and empirical data. Finally, I will cite the criticisms that have been made of agent-based simulation and the responses that might be made to these. This general introduction should also aid comprehension of the contributions published in this special issue. For readers who want to pursue the subject further, in addition to a long bibliography, I have also provided suggestions for learning more about agent-based simulation¹ (see the “To go further” Box on pages 461–462).

1. In order to make reading easier, throughout the text I will use “agent-based simulation” as a linguistic shortcut even though it will be clear in what follows that theoretical modelling and formalization will always be prior to the simulation itself. In the same spirit I will omit the subtle distinction sometimes used in the literature between “agent-based modelling” and “multi-agent systems,” the latter being a variant of the former in which there are few “agents” that are endowed with a complex internal architecture (see, for example, Axtell, 2001: 33). In fact, the distinction relates to different disciplinary traditions (notably, on the one hand, computer science, artificial intelligence, and computer engineering, and on the other, the human and social sciences) that use the same methods but with different points of view and objectives (on this point, see also Livet, Phan and Sanders in this issue).

What is agent-based simulation?

According to a common distinction,² the term computer simulation relates to two broad classes of application: 1) the techniques used when a statistical or mathematical model raises estimation, analysis or visualization problems for which an analytical approach would be inadequate; 2) the methods used when a model is constructed with the primary aim of imitating the details of a mechanism or real process (for which it may be impossible to formulate equations in practice; on this point, see Axtell, 2000).

Although this distinction is exaggerated—a statistical or mathematical model can indeed be formulated to describe a mechanism or a real process—it draws attention to a fundamental dimension that can guide us through the variety of simulation methods: that is the theoretical richness and the realism of the model that a particular method helps to convey through computer algorithms. The enormous flexibility made possible by agent-based simulation in this procedure is the feature that I intend here to characterize as its originality in comparison to other methods of computer simulation conceived as a support for statistical and mathematical analysis, and compared to other forms of simulation also designed for a more direct imitation of real mechanisms (for an overview of the latter, see Gilbert and Troitzsch, 2005).³

To clarify this subtle point, it is instructive to compare pioneering applications of simulation in the social sciences, in particular, Raymond Boudon's (1973) and Thomas C. Schelling's (1971) analyses that respectively proposed generative mechanisms for educational inequalities and spatial segregations.

Let us begin with Boudon (1973: 108–16). To account for differences in the distribution of levels of educational achievement between social groups, the French

2. Among the philosophers of science, see, among others, Julian Reiss (2011: 244–5) and Franck Varenne (2013: § 2); among the practitioners, see, among others, Averill Law (2007: ch. 1), and, formerly, André Davidovitch and Raymond Boudon (1964: 222, 232–4) and Boudon (1965: 3).

3. The following examples may help the reader to understand what is meant by computer simulation serving as a support for statistical and mathematical analysis: a (pseudo-) random distribution of numbers can be simulated by iterating a deterministic function (Tjims, 2012: 50–5); the theoretical distribution of a certain type of random variable can be simulated by transforming a variable whose values have the same probability of appearing (Ross, 2006: § 4.1 and 5.1); the independence between two discrete variables can be simulated by constructing a fictional cross-table resulting from simple multiplication of marginal distributions (Powers and Xie, 2000: 90–3); the variability of a measure of interest can be simulated by extracting a number of sub-samples of a sample and recalculating the measure of interest for each of them (Davison and Hinkley,

1997); the probability of a given choice (as a function of a series of observed factors) can be simulated by calculating the average probability produced by this choice when random values are attributed to the error terms that synthesize unobserved factors affecting the choice in question (Train, 2009: 3–7); a purely random network (i.e., one in which each node has the same probability of being linked to each of the others) can be simulated to obtain a benchmark network for an observed network (Jackson, 2008: 9); a series of discrete time units can be simulated to approximate the solution to a differential equation (Feldman, 2012: 310–1, 323–5); the behaviour of an algebraic solution whose form is not transparent can be simulated by attributing a series of specific values to the variables that define it (Gould, 2002: 1157); the sequence of states a system can find itself in can be represented by constructing transition matrices whose probabilities only take into account the last state of the system for each time unit and, by varying these probabilities the system's most probable equilibrium state can be simulated (Tjims, 2012, chs. 15 and 16).

sociologist postulated that the actors had varying degrees of success because of the unequal distribution of cultural resources in the cultural space and that given this success, the assessment of benefits (in terms of maintaining social status) and the cost and risks of educational investment also varies as a function of social milieu. To deduce the consequences of his microscopic hypotheses, Boudon proceeds as follows: a) he creates a fictional distribution of educational achievement as a function of social group; b) he creates a fictional distribution of the probability of survival as a function of social group and achievement at each point of bifurcation in the school system; c) he determines the probability of obtaining a given educational level for a given social group by multiplying these two distributions (*ibid.*: 170–3). Through this procedure, and its extension in the case where the probabilities of survival change over time, Boudon generates the social stratifications of fictitious qualifications, the qualitative structure of which resembles that of real distributions.

Conversely, how does Schelling (1971) proceed? To demonstrate that residential segregation can occur at a systemic level even though the ethnic preferences of actors taken separately are not discriminatory, the American economist constructed a series of models whose basic principle is that the decision of a given entity to remain in the place they find themselves rather than moving depends on the proportion of similar entities they are surrounded by. Schelling thus implicitly introduces the notion of the local “neighborhood”, which would become central to the agent-based approach. In a series of experiments in which he manipulates the spatial distribution of the virtual entities, the structure of these entities’ neighbourhoods and their preferences for similar neighbours, Schelling demonstrates that the repetition of the “stay/depart” choice over time changes the composition of each entity’s neighborhood and that, at the collective level, this dynamic can lead to a spatial concentration of entities depending on their group.

A forgotten pioneer

From a historical point of view, we should note that the Swedish geographer, Thorsten Hägerstrand (1965) devised and studied an agent-based simulation even before Schelling, without naming it as such, in his study of the diffusion of two agricultural innovations in Sweden. To study the hypothesis according to which the spatial structures of this diffusion depend on the way in which information is spread through social networks (and in part on the degree of the actors’ resistance to the innovation), he represented space in the form of a bidimensional grid in which each cell is inhabited by a certain number of “robots” (*ibid.*: 50). Then, he postulated that the innovation is diffused when a “robot” that has adopted the innovation speaks to a “robot” that has not done so. According to his model, the probability that this exchange happens depends on the distance between the two cells in which the “robots” reside (Hägerstrand uses migration data to calibrate this probability). By iterating the selection of discussion partners on the basis of this simple rule, Hägerstrand produced forms of spatial concentration that resembled the real Swedish distribution. In contrast to Schelling’s model, which has been repeatedly replicated and refined (see, among others, Bruch and Mare, 2006; Clark and Fossett, 2008), Hägerstrand’s study has remained confined to the study of the diffusion of innovations (Rogers, 2003: 90). Its importance is, however, significant since it illustrates the possibility of relating an agent-based simulation to real data, while Schelling’s model is completely disconnected from any empirical data. As we will see, these two directions continue to shape the research field (on this point see also Livet, Phan and Sanders in this issue).

For our discussion, it is important to recognize the following difference between Boudon's and Schelling's approaches: while the French sociologist, despite a formulation of the theoretical model at the microscopic level, formulates a mathematical model operating on defined probability distributions at the group level and adopts simulation to study this aggregated model, Schelling directly manipulates computer entities representing actors and moves them iteratively as one would do with pawns on a chessboard (on Schelling, see also Livet, Phan and Sanders in this issue).⁴

This is the distinctive feature of agent-based simulation. It helps avoid a gap being created between the level of abstraction at which hypotheses are made and that of the formal model used to study them. If there is a transition from a "low" to a "high" level, this method enables it to be recreated. Moreover, as suggested by Schelling's example, since we are not obliged to use a simplifying mathematical model, agent-based simulation enables us to remove as many theoretical simplifications as we wish.⁵

What is agent-based simulation more precisely? An "agent", in a computational sense, should be thought of as a computer program in itself (Wooldridge, 2009: 5). This program can be simple (as in the cases of Schelling and Hägerstrand) or complicated (as in a robot capable of interacting with a human being). The presence of at least two agents creates a "multi-agent system", defined by Michael Wooldridge as a system "that consists of a number of agents, which interact with one another, typically by exchanging messages through some computer network infrastructure." The system is far from operating as a homogenous whole. Quite the contrary, a) each agent can have its own attributes and its own states; b) each agent can be designed as driven by rules that are its own; c) each agent can be inserted into a geographical or relational space that limits its behaviour; d) the behaviour of each agent can depend on the behaviour of other agents in its local space; e) each agent possesses variable quantities of information (Epstein, 2006: 6). In the context of a multi-agent system, simulating means asking each agent repeatedly to execute the rules that define them. In the course of these iterations, the aggregated results of agents' behaviour can be

4. Fararo and Kosaka (1976) had no difficulty reconsidering Boudon's model by completely excluding his microscopic verbal formulation and then proposing a version in continuous time through use of differential equations. The approach Schelling followed is, in contrast, now recognized as an *ante litteram* agent-based simulation (on this point, see also the new material discussed in Hegselmann, 2012).

5. Marc Granovetter's (1978) pioneering study also illustrated this point. His objective was to study the situations where aggregation of actors' preferences is not sufficient to account for the collective result because of the interdependence between the actors' choices. Granovetter focuses in particular on binary choices and imagines that each actor chooses option A rather than B on condition that a number of actors have already opted for option A. Granovetter calls this number the "threshold" (*ibid.*: 1422). The model is used to determine

how many actors engage in a given action and at which point the number of engaged actors stabilizes for a given distribution of actors' "thresholds". Yet, despite the verbal microscopic formulation of a model, Granovetter formalizes it by use of a difference equation (which, starting from the frequency distribution of thresholds, gives, for each unit of time, the number of actors who can be engaged as a function of the number of actors engaged in the previous time unit) and he studies this equation using a combination of numerical examples and graphics (*ibid.*: 1425–8). On several occasions, the American sociologist also notes that his equation does not enable him to study finer variations of the model taking into account, for example, the hypothesis that active actors could have varying influences on a non-active actor's choices depending on a friendly relationship between them (*ibid.*: 1429–30). To do this he admits that computer simulation would be necessary.

determined step by step and be reinjected into the behaviour of these same agents. Thus, through a dynamic chain of loops connecting different levels of abstraction, agent-based simulation enables the behaviour of “low” level entities to be combined to generate the macroscopic regularity that we want to reproduce.

Agent-based simulation thus allows us to go beyond the forms of numerical simulation adopted by Boudon and Granovetter, which in fact used simulation in support of a mathematical model formulated at the aggregate level. It also enables the generation, in the sense of producing gradually “from below”, of distributions and sequences between a system’s states that Boudon and Granovetter were obliged in contrast to postulate. To express this difference, Michael Macy and Andreas Flache (2009: 248) suggested thinking of agent-based simulation as a method in which “a single unified model of the population” is replaced with “a population of models each of which is an autonomous decision-maker”. This decomposition makes it possible to model mechanisms governing the dynamics of each level of analysis and their relationships.

To conclude, let us take a final step in the characterization of agent-based simulation. Very often this method is presented *solely* as being *qualitatively* different to forms of simulation based on mathematical equations and, in particular, on equations linking macroscopic states (see, for example, Macy and Flache, 2009: 251–61). As we have just seen, this representation is not incorrect. It is, however, a partial one. On the one hand, as highlighted by Epstein (2006: 27, 54–6), it tends to create a false opposition between simulation and mathematics, while, in principle at least, an agent-based simulation can always be reformulated in terms of equations (on this point, see Treuil, Drogoul and Zucker, 2008). On the other, the fact that agent-based simulation frees the researcher from constraints imposed by other forms of numerical simulation does not mean that the latter are incompatible with agent-based simulation. The opposite is true. As the application of this method in biology shows (for two examples, see Zhang *et al.*, 2009; Wang *et al.*, 2013), various current mathematical and numerical techniques can coexist within an agent-based simulation.

This is why I suggest that, in comparison to mathematical modelling based on aggregated variables, to types of numerical simulation in support of mathematical models with no analytical solutions, and to other forms of computer simulation closer to agent-based simulation—such as cellular automata (see, among others, Feldman, 2012, chap. 27)—the best way to characterize the novelty of agent-based simulation is to stress its flexibility. That is, its flexibility from the point of view of the details and levels of analysis that it enables to be represented and to be related to each other, but also its flexibility in terms of the formalism that can coexist within the same agent-based simulation (on this point, see also Varenne, 2011: 171–3, 2013: § II.4.3).

Types of explanation and agent-based simulation

We are so used to associating the practice of simulation with the use of computers that it may be surprising to learn that Schelling studied his *ante-litteram* agent-based simulation by hand (Hegselmann, 2012). As the economist Joshua M. Epstein maintains, “[...] the essential move is conceptual, not technological. [...] *The computer is not the point*” (2006: XIII, emphasis added). It is this idea that I would now like to develop. Beyond its flexibility as a modelling method, agent-based simulation should in fact be appreciated in sociology for the type of explanation it allows.

J. M. Epstein calls this type of explanation “generative.” By this he means that the regularity of interest is generated or reconstructed from the “bottom” in numerical form through the step-by-step repetition of a series of rules governing the behaviour and interactions of entities that are assumed to have participated in the production of the regularity in real life. According to Epstein, the generative conception of explanation contains the following imperative: “If you did not grow it, you did not explain it” (*ibid.*: 67). He considers that the essence of agent-based simulation is precisely this generative capacity in the sense that an agent-based simulation “[...] provide(s) computational demonstrations that a given microspecification is in fact sufficient to generate a macrostructure of interest “ (*ibid.*: 8).⁶ Although formulated differently, these ideas are not alien to our discipline. In an article on the way in which statisticians have influenced the concept of causal explanation in sociology, the British sociologist John Goldthorpe (2001) distinguishes three conceptions of causation. According to the first, which he calls “causation as robust dependence”, a causal explanation amounts to proving that a zero-order correlation between two variables is resistant to the introduction of third variables (*ibid.*: 2–4). He then identifies a form of “causation as consequential manipulation”, according to which a causal explanation amounts to modifying the values of a given fact and to proving that a modification in the variable of interest follows from this (*ibid.*: 4–8). Finally, he identifies a conception of causation that he labels “causation as generative process,” according to which a causal explanation can only be given when we formulate hypotheses on a “process” at a lower level than the data to be explained, and that we prove that this process “does in fact operate to produce, or help to produce” this data (*ibid.*: 12).

Goldthorpe justifiably raises the problem of the method to use to furnish this proof. Indeed, when we conceive of causation as robust dependence or as consequential manipulation, to neutralize the effect of hidden variables that can add “noise” to the effect of *X* on *Y*, there are multivariate statistics and the experimental method (or if the latter is not applicable, sophisticated variants of multivariate statistics that attempt to approach experimental conditions based on observational data; see Winship and Morgan, 1999). In contrast, when it comes to testing the existence of an unobservable process underlying the data, Goldthorpe admits that the methodological options are limited. He proposes two (*ibid.*: 13–14). The first is an indirect strategy deriving the postulated process from consequences that are not directly linked with the regularity that it is the intention to explain, and demonstrating the existence of these “secondary” consequences using traditional statistical tools. The second is a direct strategy formulating an explicit (statistical) model of the process that is thought to be at work and studying it by simulating it.

Although Goldthorpe concludes by making it clear that quantitative sociologists have not yet greatly developed simulation, what should be emphasized here is that coming from different scientific traditions and disciplines Epstein and Goldthorpe make this association between a generative conception of explanation and analysis of a theoretical model of a mechanism through computer simulation. Goldthorpe does not refer specifically to agent-based simulation but the basic argument is the same.

6. This is a “demonstration”, notes Epstein, since a simulation is based on a strict process of deduction. In fact, the execution of a computer program containing rules of behaviour and interaction amounts to the deduction of the consequences of these rules, given the numerical starting conditions (i.e., the simulation’s inputs). This is why, he continues, while it is true that simulation does not lead to theorems, simulation shares its deductive nature with mathematical models.

Today this argument can be found in the sociological approach called “analytical sociology” (Hedström, 2005). Since the 1990s, this perspective has sought to develop older intuitions concerning the role the concept of mechanism could play in the construction of precise and detailed explanations in sociology (for a review of the historical development of these ideas, see Manzo, 2010). The central idea is that an explanation gains depth as we hypothesize on the way in which a set of entities interact to create a link between two (or more) events, as well as on the properties and activities characterizing these entities. Following a definition from biology (Machamer, Darder and Craver, 2000), analytical sociology thus defines a mechanism as a triad of “entities/properties/activities,” while the specific form of the elements of this triad depend on the regularity being examined. Insofar as a mechanism so defined is not, in general, directly observable, this means constructing a “model”, i.e., a simplified theoretical representation. Since the epistemic function analytical sociology assigns to a model of a mechanism is that of accounting for the genesis of an association between two (or more) events, on the basis of older definitions (Fararo, 1969) I propose calling this specific form of model “generative models” (Manzo, 2007). What I am interested in highlighting here is that analytical sociology recognized that agent-based simulation and modelling is a powerful tool for formalizing and studying generative models (Hedström, 2005, ch. 6; Hedström and Bearman, 2009: 16; Manzo, 2014). Here we thus find the association between computer simulation and a “generative” conception of explanation independently proposed by Epstein and Goldthorpe (in political science, see Cederman, 2005).

To conclude, it is not without significance that this generative and “reconstructive” conception of explanation is at the root of so-called “computational” biology. A growing number of studies on cellular growth, the development of vessels, and the immune system, are indeed based on the ideal that an understanding of these phenomena can progress provided that they can be reconstructed *in silico* and by programming the elementary mechanisms. On the methodological level it is in effect agent-based simulation that is proposed as the most flexible tool to carry out this reconstruction (for two examples of this literature, see Thorne, Bailey and Peirce, 2007; Chavali, Gianchandani, Tung *et al.*, 2008).⁷

Objects, mechanisms, and processes

To better understand the source of agent-based simulation’s flexibility and its generative ability, it is useful here to describe the technical infrastructure underlying this method. To do this, the concept of “object” should be introduced. In the sense

7. The fact that several authors with different theoretical, methodological and disciplinary orientations find the same association suggests that we are faced with a general phenomenon that very likely relates to the type of knowledge facilitated by computer programming. The mathematician and logician, Gregory Chaitin (2006: XIII), expressed this idea in the following way: “Well, the computer changes epistemology, it changes the meaning of ‘to understand.’ To me, you understand something only if you can program it (You, not someone else!). Otherwise you don’t really understand it, you only think you understand it.” Chaitin’s claim arises from the fact that programming requires the specification of each detail in the chain of entities and events leading from a set of starting conditions to a given result, and the execution of the program allows us to test whether the “chain” of postulated details leads effectively to generating the expected result.

used in computer science, “objects” can be defined as “computational entities that encapsulate some state, are able to perform actions, or methods, on this state, and communicate by message passing” (Wooldridge, 2009: 28). The similarity between this definition and that of “agent” and “multi-agent system” introduced in the first section stems from the fact that “objects” are the “deep” units of a multi-agent system, that is, the basic elements that are programmed when these types of model are constructed (Hummon and Fararo, 1995). A multi-agent system can in fact be seen as a set of object classes, a “class” being a set of objects that share the same properties (or “attributes”) and the same rules (or “functions” or “methods”, depending on the terminology of the programming language being used).

Conceiving of a multi-agent system in terms of objects enables a better understanding of the root of the distinctive characteristics I ascribed to agent-based simulation in the two previous sections, namely the flexibility of the method and its ability to provide explanations in terms of mechanisms. Let us now return to these two points in the light of the notion of the object.⁸

In terms of the flexibility of multi-agent systems, the concept of the object firstly allows us to observe that the basic unit of a multi-agent system is conceptually empty. Depending on the set of attributes and rules the modeller/programmer decides to assign to a given class of objects, these objects can represent particles, cells, individuals, groups, organizations or spatial entities (such as buildings or roads), for example. Contrary to what the term “agent” suggests, the concept of an object thus enables us to see that anything can be programmed in terms of objects, provided that we have a clear theoretical understanding of what we want to represent. A multi-agent system is therefore not limited to representing individuals. It is as much from its foundations in terms of objects that the method derives its flexibility in the representation of four central elements of sociological analysis: a) the logics of individual action; b) heterogeneity; c) interdependence between actors; and d) multiple levels of analysis.

Individual action

In terms of action theory, the “empty” nature of an object implies that any sort of rule can be defined to characterize the behaviour of objects belonging to a given class. The objects can certainly be programmed to represent actors accomplishing more or less sophisticated utility calculations (see, for example, Shoham and Leyton-Brown, 2009). However, contrary to what one often reads (see, for example, Elster, 2009: § 2), there does not necessarily have to be a link between agent-based simulation and rational choice theory (a point well illustrated by Noguera *et al.* and by Squazzoni’s review article in this issue). Objects can be used to construct agents’ internal architectures thanks to which emotions and beliefs interact in a complex way without necessarily leading to utility maximization calculations (on “cognitive agents”, see Wooldridge, 2000). Multi-agent systems are also the realm of “heuristics” in the sense

8. A possible objection to the relationship I am trying to establish between the flexibility of a multi-agent system and object-oriented programming is that, in principle, a multi-agent system can be constructed using classical procedural or functional languages (see, on this point, Izquierzo, Izquierzo and Segismundos, 2013, § 3). The response to this objection is that, in practice, the more complex a model becomes, the greater the effort needed in terms of programming when not using a tool that either directly or indirectly uses object-oriented programming.

that objects can be designed to execute very simple direct rules of behaviour (Miller and Page, 2004: 10). It is moreover to multi-agent systems that economists turn when they need to study certain “games” in which actors are supposed to follow simple behaviours iteratively, such as imitating the behaviour of the actor who was most recently successful (Alexander, 2007: 38–42; Gintis, 2009: 72–3).

Heterogeneity

Objects enable us easily to model at least five forms of heterogeneity. Within the same class of objects, the objects can have different values for the attributes belonging to this class. Between classes, objects in contrast are by definition heterogeneous in terms of attributes. Objects can also be heterogeneous in terms of the temporal sequence in which the actions assigned to them have to be executed. Finally, classes of objects can represent entities relating to various levels of analysis. The heterogeneity of beliefs, preferences and logics of action can be represented and the macroscopic consequences of this heterogeneity can be deduced. Within a multi-agent system, the researcher is thus not constrained by the simplification contained in the “representative agent” metaphor (Gallegati and Kirman, 1999).

Interdependence

The fact that the objects can “communicate”, according to Wooldridge’s definition introduced earlier, enables an understanding of the deep root of the ability of multi-agent systems to represent the interdependence of the entities of interest. The fundamental procedure through which this is achieved is to enter (numerically or logically) the values of a property of object *A* into one of the rules responsible for the behaviour of object *B*. By creating a transfer between objects in the computer’s memory, we can then define the pertinent entities for a given entity (the concept of “neighbourhood” to which I referred when describing Schelling’s pioneering model). Thus a multi-agent system makes it possible to create structures that are spatial (in which the physical distance between entities defines the neighbourhood) and relational (in which a link between *A* and *B* defines neighbourhood) within which entities are supposed to evolve (spatial and relational neighbourhoods can also be interlinked). It thus becomes possible not only to represent any sort of interaction structure, but also to make the behaviour of entities (and their internal states) dependent on the state of their local neighbourhoods as well as the global structure of interactions.

Levels of analysis

Finally, paying attention to the empty nature of the fundamental unit of a multi-agent system, *i.e.*, the objects, provides an understanding of the root of this method’s ability to represent multiple entities representing different levels of abstraction within the same model. For example, one class of objects could represent a set of organizations while a second class of objects could represent actors; communication between the objects could then be used to represent the effects of the influence of organizations over actors (for an example of this type of complex architecture combining several levels of analysis, see Ferber, Gutknecht and Michel, 2004; Ferber,

Michel and Baez, 2005; see also the discussion of “agent-cities” in Livet, Phan and Sanders in this issue).⁹

We now have all the elements needed to understand why the fundamental unit on which a multi-agent system is constructed, namely the object, makes this method a favoured approach for studying explanations conceived in terms of mechanisms. First of all, when we think in terms of objects, a profound affinity appears between a (set of) mechanism(s) and a multi-agent system. As I defined it in the previous section, a mechanism is a set of entities, with their properties and activities that interact to produce a given result at a higher level of abstraction. We now know that a multi-agent system is a collection of objects defined by their attributes and operating rules that exchange information influencing each other. In terms of its contents, a mechanism, just as a multi-agent system, is in principle neutral. Depending on the observed phenomenon, the entities in a mechanism will be a certain type, just as the objects of a multi-agent system will take a particular form depending on the entities/actions we want to model. Thus, there is a strong structural homology in form (entities/properties/activities/interactions versus objects/attributes/functions/communications) and in the content (empty in both cases) between what we want to study—mechanisms—and the method used—multi-agent systems. Provided you have a theoretical representation of these mechanisms, the method can model their structure with any level of detail.

Note, finally, that explanation by mechanisms does not only require the precise description of the behaviour of entities and the connections between them that might have led to the facts of interest. This conception of explanation also requires that we prove, according to Epstein’s expression, the “generative sufficiency” of this description, in the sense that the fact of interest must be generated by the model of the postulated mechanism. The algorithmic nature of a multi-agent system makes this reverse engineering possible.¹⁰ In effect, a multi-agent system simulation involves the step-by-step repetition of rules defining the behaviour of objects, the systematic updating of their attributes, and the exchange of information between them. In this way, the processes potentially contained in the mechanism represented by the agent-based model can run and, ultimately, the consequences of the model can be deduced and quantified. The requirement of “generativity” specific to an explanation through mechanisms is thus implemented in the dynamic part of a multi-agent system through what Patrick Doreian (1999: 98–9) called “algorithmic causality”.

There are thus profound reasons why the association between the study of mechanisms and agent-based simulation is often made in sociology, as elsewhere (in biology, see, for example, Thorne, Bailey and Peirce, 2007: 252), which relate to the internal functioning of the method. The “object”-based structure enables the most flexibility in the representation of the details of a mechanism while the step by step updating and the communication between these objects ensure that dynamic consequences, *i.e.*, the process of the mechanism, can be activated. What therefore becomes possible, ultimately, is the determination of macroscopic consequences of the (model of a) mechanism under observation, whatever its form or complexity.

9. I am aware that objects are all located at the same level of analysis, in the sense that they are all stored in the computer’s RAM (Boschetti, 2011). The issue is that this physical reality does not imply that objects cannot be designed and arranged (through the flow of information transmitted between them) in such

a way that they represent entities relating to different levels of analysis (Epstein, 2014: 165; for an example in biology, see Wang *et al.*, 2013).

10. In general, reverse engineering involves trying to reproduce the behaviour of an object (software for example) without knowing its inner workings (Eilam, 2005).

The micro-macro transition problem

The discussion of the “internal” structure of a multi-agent system that I offered above can be used in support of the idea that agent-based simulation opens promising avenues for tackling a central issue in sociology, the “transformation problem”, or alternatively the “micro-macro” transition (see Wippler and Lindenberg, 1987 and Cherkaoui, 2003 respectively); that is, the analysis of mechanisms that enable the behaviour of low-level entities to combine to produce regularities at a higher level of abstraction.

In a classic article, James S. Coleman (1986: 1321) recognized the centrality of the transformation problem and went as far as to assert that this problem, and not that of the refinement of action theory (Coleman was then in debate with Talcott Parsons), is “the major theoretical obstacle” to the development of sociological theory. What matters for my purposes here is the way in which Coleman justifies this assertion. According to him, the transformation problem is underdeveloped because while quantitative techniques are well equipped to study the effects of belonging to a particular group or a particular behaviour, sociologists do not possess comparable refined techniques for studying systems of interdependent entities and the way in which these interdependencies combine in a dynamic way to produce macroscopic effects (*ibid.*: 1316, 1329).

When Coleman was developing this argument, although some embryonic forms of agent-based simulation, such as Schelling’s, had already been formulated, it was difficult to imagine the method’s recent developments. The preceding discussion suggests that today multi-agent simulation is a powerful tool for addressing the “transformation problem” because of the way in which it enables, on the one hand, the modelling of low-level entities’ behaviour and, on the other, how it can tackle the interdependence between these entities.

Regarding the first point, in economics, as in some sociology, the strategy that is frequently adopted to deal with the transformation problem has been to give the actor a very simple image (as a device carrying out optimization calculations) and to suppose that all actors follow the same logic of action, which, because of the postulated homogeneity, enables all actors to be treated as if they were one and thus to apply the “representative agent” shortcut (on this point, see Alan Kirman, 1992). The argument that the image of the actor should remain simple in order to avoid the aggregation problem becoming too difficult to examine (mathematically in particular) is, moreover, often found within a particular actionist inspired tradition (Coleman, 1990: 19; for a more recent discussion, see Raub, Buskens and van Hassen, 2011: § 5).¹¹

Agent-based simulation provides us with a way out of this impasse because, since the base entities are individually modelled in the form of “objects” and the result of their behaviour at a given aggregation level is determined step-by-step, neither the heterogeneity nor the type of actions that we grant to these entities is an obstacle to

11. Note that the realism that is the goal of an agent-based model can result in a more complex image of the actor just as much as it can result in a greater simplification. In his latest book, Epstein (2014: 3) draws inspiration from the neurosciences to increase the inner complexity of agents and remarks that, while this discipline looks for more complexity within the actor, it does not have the tools to describe the collective consequences of this infra-individual complexity. Agent-based simulation, he emphasizes, should therefore be urged to do something about this. Many recent

studying the transition from the “micro” to the “macro” level. Agent-based simulation thus overcomes this compromise between simplistic microscopic hypotheses and the feasibility of establishing macroscopic consequences by which a large part of mathematical modelling and natural language are constrained.

The interdependency of actors, the aspect which, according to Coleman, is the principal obstacle to modelling the passage from the microscopic to macroscopic level, is the second element that can be particularly flexibly handled by agent-based simulation. We have seen that one of the typical features of “objects” is that they are able to communicate between themselves and to be arranged in spatial and/or reticular structures such that the behaviour of each artificial agent can depend on the agents that make up its “neighbourhood”. This dependency can be modelled in a very detailed way, since by playing with the “object” attributes we can represent the agents’ inner components, such as their beliefs, reasons and emotions, and make these evolve dynamically as a function of the behaviour of the agents’ neighbourhood or even a specific agent. The fundamental point here is that, since the elementary rules of behaviour are executed step-by-step, and the result of the local influences to which these behaviours are submitted calculated progressively (by reinjecting these results in the next iteration), the macroscopic result can be derived. Essentially, it is the recursive nature of simulation that enables the determination of the macroscopic implications of the local macro–micro influences embedded in a given interdependent structure.

To conclude, let us ask why agent-based simulation would be better able to tackle interdependence in the micro–macro transition than social network analysis methods (methods that Coleman himself considered to be a promising avenue) (the article by Gabbriellini, in this issue, also addresses this question). If we consider the most sophisticated statistical modelling of relational data (for an overview see Tom A. B. Snijders, 2011), the difference between it and agent-based simulation *stricto sensu* is two-fold. On the one hand, on the downside of effects constraining actors’ behaviour, the postulated mechanisms in these statistical models are limited to structural properties, such as transitivity, without illuminating the microscopic origins of these effects. An agent-based model can, in contrast, describe how each of these properties has an effect on agents’ behaviour, by for example clarifying how the topology of the network impacts on the components of the actors’ actions, such as opportunities and/or beliefs. On the upside, then, the most sophisticated forms of statistical analysis of social networks are intended to calculate the probability of the creation/destruction of a tie and, consequently, the only types of macroscopic consequences these statistical models are interested in is the structure of the network itself. Agent-based simulation, in contrast, can insert interdependencies mediated by a network of ties into a more general mechanism and is thus interested in the macroscopic consequences situated at a higher level of abstraction than the network itself.¹²

(note 11 continued)

studies in cognitive psychology in contrast, have disregarded the inner complexity of actors and experimentally study the shortcuts (the “heuristics”) that the latter use to manage the complexity of social life (for an overview of this literature, see Gigerenzer and Gaissmaier, 2011). Agent-based simulation can also be used to derive the macroscopic consequences of models postulating networks of actors following a logic of *homo heuristicus* (for two

examples, see Todd, Billari and Simao, 2005; Manzo and Baldassarri, 2015).

12. It is indeed to agent-based simulation as a complement to network analysis that we turn when the objective is to model the mechanisms responsible for a particular structural property in a more detailed way (Monge and Contractor, 2003: ch. 4) and when the aim of the analysis is to study the systemic consequences of a particular network structure (see, for example, Centola and Macy, 2007).

Agent-based simulation and empirical data

So far I have presented agent-based simulation as being capable, as a result of its deep structure, of implementing theoretical models with whatever detail and of using these to reconstruct *in silico*, in numerical form, the regularities we want to explain. Conversely, I have not mentioned the relationship that can exist between agent-based simulation and empirical data.

According to some, this is not a big omission. Indeed, agent-based simulation (and simulation in general) is often considered to be incapable of making inferences about the real world (see, among others, Grüne-Yanoff, 2009). According to this point of view, this method can only be an instrument for theorizing. In light of the variety of strategies that are available today to relate agent-based simulation to empirical data and to assess the robustness of a simulation, this criticism can be considered to be based on a simplistic vision of the method (for analyses along these lines, see Muldoon, 2007; Reiss, 2011: 253–4; Casini, 2014). It is thus to these strategies that relate agent-based simulation to empirical data that I would briefly like to turn.¹³

Depending on available data, one possibility is to compare the macroscopic numerical structures (and dynamics) generated by the model simulation to aggregated quantitative data from surveys, administrative sources or digital traces of various types (such as those left on the internet) enabling the objectivization of macroscopic regularities of interest. To make this comparison, statistical analysis can be used both to describe the simulated and empirical data in parallel and to assess the differences between the two data sets (see Gabbriellini in this issue). This process is not simple (Fagiolo, Windrum and Moneta, 2007) and new procedures are regularly proposed (Thorngate and Edmonds, 2013). It can nevertheless be done rigorously when the data allow, either from a transversal or diachronic perspective (for two examples, see Manzo, 2013a and Gonzalez-Bailon and Murphy, 2013, respectively).

The second agent-based simulation validation strategy is linked to the limitations of the first. Indeed, the congruence between simulated and real macroscopic structures is not proof in itself of the realism of the microscopic and relational details designed to generate the macroscopic structures. In Epstein's (2006: 8–9) terminology, this congruence only proves the model's "generative sufficiency", but it does not exclude the fact that (partially) alternative microscopic specifications could generate equally realistic macroscopic structures (and dynamics). Empirical "calibration" of an agent-based simulation is a strategy destined to respond to this problem. The

13. The position I am defending here takes account of the point made by Michael Macy and Yoshimichi Sato (2008: 1.4) when they note that "it is empirical research, not an agent-based model, that bears the burden of proof. The computational model can generate hypotheses for empirical testing, but it cannot 'bear the burden of proof'." Moreover, it is incontestable that wanting to relate an agent-based simulation to empirical data has its dangers. Indeed, if the behaviour of the model has not previously been studied systematically over its entire parameter space (I will come back to this point in the following section), the researcher may fail to identify strange results (that could result, for example, from a programming error) or incoherent results in relation to the theoretical signification attributed to the represented mechanisms (signalling that these mechanisms have been incorrectly formalized). Thus, approaching empirical data too early can lead to wanting to validate a model that does not (yet) deserve this effort. While a purely theoretical exploration of an agent-based model thus has a value in itself, it is completely legitimate (and technically possible) to connect an agent-based simulation to empirical data once we have sufficient confidence in its solidity and theoretical coherence.

general idea is to base the microscopic and relational specification of the simulation on empirical data so that the realism of the “low” level specification can be justified.

This calibration can be done indirectly or directly. In the first case, qualitative and ethnographic observations (see Casilli, Rouchier and Tubaro in this issue; more generally, Moss and Edmonds, 2005; Moss, 2008) or the results of laboratory experiments (Duffy, 2006; Fischbacher and Gächter, 2010; Boero *et al.*, 2010) are used to define the agents’ behaviour and interactions. When calibrating an agent-based simulation directly, in contrast, empirical information is introduced to several of the model’s components in a quantitative form (as Noguera *et al.* and Gabbriellini do in this issue). Depending on the model studied and the available data, we can for example, input the values of the agents’ attributes based on the real probability distributions of these attributes or randomly extract these values from probability distributions adjusted to real distributions (Brown and Robinson, 2006; Bruch and Mare, 2006). Functional (probabilistic) forms estimated from real data can also be used to represent the way in which the artificial agents’ attributes combine and lead them to a particular behaviour (Hedström, 2005: ch. 6). In a similar way, information from empirical networks can be used to define the properties and topology of an artificial network in which agents are supposed to evolve (Rolfe, 2014). Using georeferenced data, we can even define the spatial structure of agents’ interactions on the basis of real physical distances (Crooks and Castle, 2011; Girardin and Cederman, 2007).

By combining these different dimensions, when possible, we quickly arrive at agent-based simulations that are tightly constrained by empirical data and have a high degree of realism, in terms of the initial conditions for the agents’ behaviour, their interactions, or their environment. Some now classical studies in archaeology (for an overview, see Epstein, 2006: chs. 4–6), and epidemiology (for a recent example, see Frias-Martinez, Williamson and Frias-Martinez, 2011) amply illustrate these developed forms of empirical calibration, the objective of which becomes the almost perfect reproduction *in silico* of the phenomenon of interest and the prediction of its future states rather than the elimination of competing micro- and meso-specifications.¹⁴

The connection between agent-based simulation and empirical data is undoubtedly essential for this method to achieve the explanatory objectives attributed to it. However, one must be aware that asking too much from empirical calibration, at the “macro” as well as “micro” level, can lead the researcher to select the aspects to model on the basis of the availability of empirical data. In this sense, empirical calibration contains within it a “conservative” tendency at the theoretical level (on

14. When this variant of empirical calibration is used, agent-based simulation approaches other forms of computer simulation, such as micro-simulation, which also model the behaviour of each entity singularly on the basis of random empirically-calibrated events (for a detailed analysis of micro-simulation in demography, see Imhoff and Post, 1998). Agent-based simulation differs, however, from micro-simulation because it always leaves open the possibility of detailing something for which we have no empirical data in order to estimate a probability that we nevertheless want to explore theoretically (Billari *et al.*, 2006: 8). Work comparing agent-based simulation and micro-simulation (see, among others, Mark Birkin and Belinda Wu, 2012; Eric Silverman *et al.*, 2013), stresses the greater flexibility of the former for modelling the effect of “high” level feedback on the “low” level as well as the complexity of behaviours and interactions (on interdependence in micro-simulation, see Imhoff and Post, 1998: 117–9). This is why Francesco C. Billari and Alexia Prskawetz (2003: 4) suggest that micro-simulation should be considered to be a special case of agent-based simulation.

this point, see Fagiolo, Windrum and Moneta, 2007: 211–2). Thus, a compromise should be found on a case by case basis between, on the one hand, the underlying reason for using agent-based simulation as opposed to other forms of simulation, namely the possibility to detail and investigate aspects for which there is insufficient empirical data, and on the other hand, the legitimate desire to constrain the model through empirical data.

Criticisms of agent-based simulation

The objection that inspired the previous section, according to which agent-based simulation can only serve as a theoretical exploration tool, lies behind other criticisms regularly made of this method (made with particular force by economists, see Leombruni and Richiardi, 2005, for an overview). Since the results of agent-based simulation depend on initial conditions (Fararo and Kosaka, 1976: 431–3; Sørensen, 1976: 85, 89; more recently, Gould, 2002: 1169–70), are uncertain (Reiss, 2011: 258–9), are not transparent (Young, 2006; Gintis, 2013; Treuil, Drogoul and Zucker, 2008, ch. 7) and are only moderately comparable and reproducible (Fagiolo, Windrum and Moneta, 2007: 198), agent-based simulation could not contribute to an understanding of the real world and should, therefore, be limited to guiding our theoretical interests.

Although it is undeniable that these are valid criticisms, when looking in detail at how agent-based simulation works in comparison to other methods that it is suggested are more reliable, such as laboratory experiments, mathematical models or statistical analysis and econometrics (for comparative analyses of this type, see, for example, Leombruni and Ricchiardi, 2005; Winsberg, 2009; Reiss, 2011), we see that agent-based simulation's detractors (and those of simulation more generally) tend to underestimate the presence of similar problems for the methods they prefer; the cost of resolving certain problems with the help of these methods; and the solutions that can be used to increase the reliability of results produced by (agent-based) simulation. Without disregarding the fact that these solutions are still improvable and difficult to implement (Thiele, Kurth and Grimm, 2014), I would like to focus on the latter in what follows.

The conditional nature of agent-based simulation's results

Simulated results always depend, on the one hand, on the parameters' values under which the simulation was performed, and, on the other hand, they depend on each detail of the "internal" structure of the model. The strategies to pursue to address the problem differ depending on available empirical information.

Dependency with respect to input values

In terms of the dependency of results on the chosen values for the model's parameters, if the researcher at least has accurate data on the macroscopic structure of interest and if he has sound theoretical and empirical reasons to believe in the microscopic specification of the model, it is possible to "fix" the value of these parameters through the use of non-parametric iterative procedures that change the parameters' values until the greatest proximity between the observed and simulated macroscopic

structures is achieved (see Gabbriellini in this issue). This form of parameterization is common among agent-based simulation's practitioners: it is often called "calibration" (Railsback and Grimm, 2011: ch. 20).¹⁵ Conversely, when the researcher does not possess sufficiently detailed empirical data to "fix" the value of the parameters, analysis of the model can only be based on what is called "sensitivity analysis". Beyond the variety of available technical options (for a detailed review, see Saltelli, Chan and Scott, 2000; see also, Stonedahl and Wilenski, 2010), this approach can be understood in two ways. In a strict sense, sensitivity analysis amounts to evaluating how variable a model's given result is for a given variation of one of its parameters (this is what is sometimes called "local sensitivity analysis," see Railsback and Grimm, 2011: 293–5; Casilli, Rouchier and Tubaro in this issue follow this procedure). In a broader sense, sensitivity analysis means any strategy aimed at studying the way in which the model's behaviour changes in its parameter space, that is across all possible combinations of values that each parameter of the model can receive (this is what we call "global sensitivity analysis," see Railsback and Grimm, 2011: 297; Noguera *et al.* and Gabbriellini use a similar procedure). Multivariate statistical analysis can then be used to describe the variations in the simulated results depending on the variations in the parameter values (for two examples, see Fararo and Butts, 1999: 51–2; Manzo and Baldassarri, 2015).

Dependency with respect to the internal structure

As for the model's results' dependence on its internal structure, when we do not have sufficiently detailed empirical information to "fix" this structure (in other words, when the forms of "calibration" discussed in the previous section are not feasible), the avenue to pursue is that of "robustness analysis" (Railsback and Grimm, 2011: 302–6). In contrast to sensitivity analysis, it is not easy here to suggest general recipes because, by definition, the internal structure of a model is specific to that model (Epstein, 2006: 30–1). As a general rule, however, assessing result stability is a matter of adjusting the following aspects: a) the form of probability distributions used for the stochastic parts of the model; b) the functional forms adopted to link the elementary entities' attributes; c) the rules of behaviour of these entities; d) the structure of their interactions (which is often done by simulating the model under several network topologies, a practice made popular by R. Axtell, 2001; see also Noguera *et al.* in this issue); e) the sequence followed to "call" each entity; and f) the way in which the behaviours of elementary entities are adapted and updated from a temporal point of view (on these last two points, see Axtell, 2001; Miller and Page, 2004).

The uncertainty of the results of an agent-based simulation

A second tricky aspect of the analysis of an agent-based simulation is the variability of results obtained depending on the internal structure of the model for the same set of parameters when the model is repeated several times (Troitzsch, 2014).

15. In macroeconomics, there is a lively debate around the possibility of calibrating theoretical models studied through simulation (see the critique by Hansen and Heckman, 1996, on the approach to the study of real economic cycles). J. Reiss (2011: 256–60) explains why it would be a mistake to transpose this debate as it stands to the field of agent-based simulation and therefore to use it to criticise this specific form of simulation.

The source of this variability is in the random elements contained in the model (such as a probabilistic decision rule determining the choice of agents, for example). It should be controlled so as not to confuse the variability due to a substantial change in the model with that derived from its random components (which can be done by controlling the “seed” that manages the generation of pseudo-random numbers in the simulation). This variability must also be quantified and described, which requires that the simulation be replicated a certain number of times for each set of parameters. This requirement increases the calculation time needed to study a model (especially when systematically examining the parameter space) and demands the researcher be creative in the presentation of results. In this regard, when looking at the macroscopic behaviour of the model, we can describe the central tendency of the distribution of repetitions accompanied by a measure of dispersion. If the objective is to compare simulated macroscopic data and real aggregated data, a possible strategy is therefore to evaluate the degree of overlap between variability in the simulation between repetitions and the empirical variability between different samples (for an example of this strategy, see Manzo, 2013a). The issue is trickier when one is interested in the trajectories of each artificial agent. In this case, the plurality of parameter combinations to study with the variability of simulations within each combination rapidly generates very large and difficult to describe data structures. Multi-level regression models have recently been suggested to deal with this twin variability (for two examples, see Manzo, 2013b; Fountain and Stovel, 2014).

The transparency of the results of an agent-based simulation

This reference to the trajectories of each agent brings us to a third thorny issue for agent-based simulation, namely understanding the inner workings of the model being simulated. By inner workings I mean the numerical details of the way in which agents evolve, their interactions and the feedback effects that take place once the simulation is launched. Indeed, although the theoretical mechanisms coded in the computer program might be perfectly clear (it is us who designed them!), the process generated by these mechanisms is not always easy to understand. Moreover, it is this process that must be understood to clarify the origin of the model's behaviour in the different areas of its parameter space.

The weight of this task varies depending on the complexity of the model's microscopic specification (Macy and Flache, 2009: 261–4). To some extent, sensitivity analysis is itself a tool to help develop intuitions on the model's workings (on this subject, see the first, second and sixth “heuristics” proposed by Railsback and Grimm, 2011: 280, 282). In a similar vein, the mathematical description of the sequences of states through which the model passes—using differential equations (Huet and Deffuant, 2008) or Markov chains (Izquierdo, Izquierdo and Segismundos, 2013), for example—can aid a better understanding of the model's dynamics. To grasp the numerical detail underlying the aggregate behaviour of an agent-based simulation, however, it is often necessary to use more specific procedures. In some cases it is possible to sequentially activate postulated mechanisms so that we can isolate their effects on both the internal dynamic of the model and its aggregated results (for an example, see Manzo, 2013a). When this operation is insufficient (or impossible because the postulated mechanisms are embedded) ad hoc measures need to be invented—which Railsback and Grimm (2011: 284) call the model's “currencies”—to study how a particular “low” level aspect evolves over time, as well as its impact on the aggregated result of interest (Noguera *et al.*, in this issue, followed this type of approach).

The reproducibility of the results of an agent-based simulation

The standardization of procedures for the construction and analysis of an agent-based model is also a critical issue in itself (for a detailed review, see Müller *et al.*, 2014). Despite the development of platforms for constructing and analysing agent-based simulations (Nikolai and Madey, 2009; Lytinen and Railsback, 2012), researchers in this field have found themselves largely in the same position as that which characterized the use of statistical methods in the social sciences until the appearance of software packages (on this point, see Gilbert and Bankes, 2002), a period in which the researcher had to write his or her own computer code implementing the required algorithms. This situation is, in part, due to the nature of agent-based simulation itself. Since it requires the formalization of mechanisms able to generate a given regularity, the type of computer code needed to implement a representation of these mechanisms depends largely on the observed phenomenon and the hypotheses one seeks to formalize.

This specificity nevertheless has a negative consequence: models have multiplied rather than accumulated (on this point, see Rouchier, 2013). Since it is not easy to explain all the details of a model in an article and the computer code is not always made available, replication of published models is difficult. When the exercise is done, analysis is generally limited to proving that a particular result of the observed model is not confirmed if a particular aspect of the model is changed (see, for example, Galan and Izquierdo, 2005; Meadows and Cliff, 2012; Van de Rijt, Siegel and Macy, 2009; Wilensky and Rand, 2007) and, often because of the lack of information provided, the analysis results in a controversy as to what the authors could replicate compared to what they actually replicated (on this point, it is, for example, instructive to read the following in order, Will and Hegselmann, 2008; Macy and Sato, 2008).

To create more transparent communication that could aid both re-use and replication of models, two trends have emerged. On the one hand, protocols for the standardized presentation of agent-based models have been proposed (Richiardi *et al.*, 2006; Grimm, Berger and DeAngelis, 2010). On the other, platforms for the classification and sharing of models—either generic (such as OpenAbm; see Janssen *et al.*, 2008) or specific to a particular language (such as NetLogo)—have developed. Although it may seem like an aside, we should note finally the growing awareness of the importance that teaching this method (still rare in university social science programmes) could have for the diffusion of good methodological practice (on this point, see Macal and North, 2013).

*

* *

In the social and human sciences, in France as elsewhere, agent-based simulation is only used by a minority of researchers. For this reason, the introduction I have written to this method is aimed more at readers curious to discover it than at specialists. My aim was to provide a guide to navigate a way through the principal theoretical and methodological debates that shape the rich and continuously expanding literature on agent-based simulation.

Thus, relying on examples, I first clarified what agent-based simulation is and in what ways it differs from other forms of computer simulation. In particular, I argued

that flexibility is a distinctive feature of this method, both from the point of view of what it allows us to represent and the formalism that it enables us to manipulate. Next, I stressed the fact that agent-based simulation enables the implementation of a specific conception of explanation, namely the idea that explaining amounts to reconstructing the explanandum, and I demonstrated that this conception was associated with computer simulation in various disciplines (including beyond the social sciences). Finally I showed the “deep” source of agent-based simulation’s flexibility, notably by delving into the details of the type of programming languages underlying the construction of a multi-agent system. This more technical section also allowed me to show why agent-based simulation opens promising avenues for the construction of theoretical models that deal with, rather than avoid, the “micro–macro” transition problem. Finally, I addressed the criticisms that agent-based simulation has to overcome in order not to descend into “science fiction”, which led me on to a discussion of strategies that can be used to relate an agent-based simulation to empirical data, in order to evaluate the stability of simulated results and to understand the inner workings of a simulation.

The foregoing analysis is an enthusiastic one, well-disposed and open to the potential of agent-based simulation. As my roundup of the criticisms of this method suggested, my introduction (and this special issue more generally) is not however intended to be a naive *pro domo* defence of agent-based simulation. Indeed, the history of the discipline invites prudence. We need only reread Boudon (1965) and Padioleau (1969) to appreciate that in the 1960s and 1970s, computer simulation was perceived as being able to provide sociology with more theoretically and empirically solid foundations. The current state of the discipline shows that simulation, however, is still not part of its current methods. Moreover, the articles by Boudon and Padioleau raised the epistemological and methodological problems I discussed, problems that evidence shows are far from being resolved. Thus, despite the undeniable power of agent-based simulation in comparison to simulation methods that have been available for around fifty years, it is completely legitimate to feel a sense of déjà-vu towards the enthusiasm generated by this method.

So, in conclusion, allow me to indicate two avenues that should be actively pursued in order to reduce the likelihood that this enthusiasm will once again prove to be misplaced. On the one hand, our ability to use agent-based simulation in conjunction with other methods could help this technique better connect to empirical data and better help us to understand the functioning of the models studied (I have stressed this point elsewhere, see Manzo, 2014). On the other hand, how we are able to integrate teaching this method in university curricula seems crucial. It is through circulation (between sub-disciplines and generations) and through teaching that these ideas will take root, spread and evolve.

Gianluca Manzo

Groupe d'Étude des Méthodes de l'Analyse Sociologique de la Sorbonne (GEMASS)
CNRS—Université Paris-Sorbonne
20, rue Berbier-du-Mets
75013 Paris

gianluca.manzo@cnrs.fr

REFERENCES

- ALEXANDER, J. M., 2007. *The Structural Evolution of Morality*. Cambridge: Cambridge University Press.
- ARVIND, K. C., GIANCHANDANI, E. P., TUNG, K. S., LAWRENCE M. B., PEIRCE S. M., PAPIN J. A., 2008. "Characterizing Emergent Properties of Immunological Systems with Multicellular Rule-Based Computational Modeling." *Trends in Immunology* 29(12): 589–99.
- AUCHINCLOSS, A. H., ROUX, A. V. D., 2008. "A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health." *American Journal of Epidemiology* 168(1): 1–8.
- AXTELL, R., 2000. "Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences." The Brookings Institution, Center on Social and Economic Dynamics, *Working Paper No. 17*. Online: www.brookings.edu.
- AXTELL, R., 2001. "Effects of Interactions Topology and Activation Regime in Several Multi-Agent Systems." In S. MOSS, P. DAVIDSSON, eds, *Multi-Agent-Based Simulation*. Berlin: Springer: 33–48.
- BILLARI, F. C., FENT, T., PRSKAWETZ, A., SCHEFFRAN, J., 2006. "Agent-Based Computational Modelling: An Introduction." In F. C. BILLARI, T. FENT, A. PRSKAWETZ, J. SCHEFFRAN (eds), *Agent-Based Computational Modelling. Applications in Demography, Social, Economic and Environmental Sciences*. Heidelberg: Physica-Verlag: 1–16.
- BILLARI, F. C., PRSKAWETZ, A., eds, 2003. *Agent-Based Computational Demography: Using Simulation to Improve our Understanding of Demographic Behaviour*. New York, NY/Heidelberg: Physica Verlag.
- BIRKIN, M., WU, B., 2012. "A Review of Microsimulation and Hybrid Agent-Based Approaches." In A. J. HEPPENSTALL, A. T. CROOKS, L. M. SEE, M. BATTY, eds, *Agent-Based Models of Geographical Systems*. Berlin: Springer: 51–68.
- BIRKS, D., TOWNSLEY, M., STEWART, A., 2012. "Generative Explanations of Crime: Using Simulation to Test Criminological Theory." *Criminology*, 50(1): 221–54.
- BOERO, R., BRAVO, G., CASTELLANI, M., SQUAZZONI, F., 2010. "Why Bother with What Others Tell You? An Experimental Data-Driven Agent-Based Model." *Journal of Artificial Societies and Social Simulation* 13(3): 6. Online: <http://jasss.soc.surrey.ac.uk/13/3/6.html>.
- BONACICH, P., LU, P., 2010. *Introduction to Mathematical Sociology*. Princeton, NJ: Princeton University Press.
- BOSCHETTI, F., 2011. "Causality, Emergence, Computation and Unreasonable Expectations." *Synthese* 181: 405–12.
- BOUDON, R., 1965. "Réflexion sur la logique des modèles simulés." *Archives Européennes de Sociologie* 6(1): 3–20. Online: <http://dx.doi.org/10.1017/S0003975600001119>.
- BOUDON, R., 1973. *L'Inégalité des chances. La mobilité sociale dans les sociétés industrielles*. Paris: Armand Colin.
- BROWN, D. G., ROBINSON, D. T., 2006. "Effects of Heterogeneity in Residential Preferences on an Agent-Based Model of Urban Sprawl." *Ecology and Society*. 11(1): 46. Online: <http://www.ecologyandsociety.org/vol11/iss1/art46/>.
- BRUCH, E., MARE, R. D., 2006. "Neighborhood Choice and Neighborhood Change." *American Journal of Sociology* 112(3): 667–709.
- CASINI, L., 2014. "Not-So-Minimal Models. Between Isolation and Imagination." *Philosophy of the Social Science* 44(5): 646–72.

- CEDERMAN, L.-E., 2005. "Computational Models of Social Forms: Advancing Generative Process Theory." *American Journal of Sociology* 110(4): 864–93.
- CENTOLA, D., MACY, M. W., 2007. "Complex Contagions and the Weakness of Long Ties." *American Journal of Sociology* 113(3): 702–34.
- CHAITIN, G., 2006. *Meta Math! The Quest for Omega*. New York, NY: Vintage Books.
- CHATTOE-BROWN, E., 2014. "Using Agent Based Modeling to Integrate Data on Attitude Change." *Sociological Research Online* 19(1): 16. 10.5153/sro.3315. Online: <http://www.socresonline.org.uk/19/1/16.html>.
- CHAVALI, A. K., GIANCHANDANI, E. P., TUNG, K. S. *et al.*, 2008. "Characterizing Emergent Properties of Immunological Systems with Multi-Cellular Rule-Based Computational Modeling." *Trends Immunol* 29(12): 589–99. Online: doi: 10.1016/j.it.2008.08.006.
- CHERKAOUI, M., 2003. "Les transitions micro–macro. Limites de la théorie du choix rationnel dans les *Foundations of Social Theory*." *Revue Française de Sociologie* 44(2): 231–54.
- CLARK, W. A. V., FOSSETT, M., 2008. "Understanding the Social Context of the Schelling Segregation Model." *PNAS* 105(11): 4109–14.
- COLEMAN, J. S., 1986. "Social Theory, Social Research and a Theory of Action." *American Journal of Sociology* 96(6): 1309–35.
- COLEMAN, J. S., 1990. *The Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- CROOKS, A. T., CASTLE, C. J. E., 2011. "The Integration of Agent-Based Modeling and Geographical Information for Geospatial Simulation." In A. J. HEPENSTALL, A. T. CROOKS, L. M. SEE, M. BATTY, eds, *Agent-Based Models of Geographical Systems*. Berlin: Springer: 219–51.
- DAVIDOVITCH, A., BOUDON, R., 1964. "Les mécanismes sociaux des abandons de poursuite judiciaires: analyse expérimentale par simulation." *L'Année Sociologique*, 3rd series: 111–244.
- DAVISON, A., HINKLEY, D., 1997. *Bootstrap Methods and their Application*. Cambridge: Cambridge University Press.
- DOREIAN, P., 1999. "Causality in Social Network Analysis." *Sociological Methods & Research* 30(1): 81–114.
- DUFFY, J., 2006. "Agent-Based Models and Human Subject Experiments." In L. TEFATSION, K. L. JUDD, eds, *Handbook of Computational Economics*, vol. 2: *Agent-Based Computational Economics*. Amsterdam: Elsevier/North-Holland.
- EDMONDS, B., MEYER, R., eds, 2013. *Simulating Social Complexity: A Handbook*. Berlin: Springer.
- EILAM, E., 2005. *Reversing: Secrets of Reverse Engineering*. Indianapolis, IN: Wiley.
- ELSTER, J., 2009. "Excessive Ambitions." *Capitalism and Society* 4: 2. Online: doi:10.2202/1932-0213.1055.
- EPSTEIN, J. M., 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press.
- EPSTEIN, J. M., 2014. *Agent_Zero: Toward Neurocognitive Foundations for Generative Social Science*. Princeton, NJ: Princeton University Press.
- FAGIOLO, G., WINDRUM, P., MONETA, A., 2007. "A Critical Guide to Empirical Validation of Agent-Based Economics Models: Methodologies, Procedures, and Open Problems." *Computational Economics* 30(3): 195–226.
- FARARO, T. J., 1969. "Stochastic Processes." *Sociological Methodology* 1: 245–60.

- FARARO, T. J., BUTTS, C. T., 1999. "Advance in Generative Structuralism: Structured Agency and Multilevel Dynamics." *Journal of Mathematical Sociology* 24(1): 1–65.
- FARARO, T. J., KOSAKA, K., 1976. "A Mathematical Analysis of Boudon's IEO Model." *Social Science Information* 15(2–3): 431–75.
- FARMER, J. D., FOLEY, D., 2009. "The Economy Needs Agent-Based Modelling." *Nature* 460: 685–6.
- FELDMAN, D. P., 2012. *Chaos and Fractals: An Elementary Introduction*. Oxford: Oxford University Press.
- FERBER, J., GUTKNECHT, O., MICHEL, F., 2004. "From Agents to Organizations: An Organizational View of Multiagent Systems." In P. GIORGINI, J. MÜLLER, J. ODELL, eds, *Agent-Oriented Software Engineering (AOSE) IV*. Berlin: Springer: 214–30.
- FERBER, J., MICHEL, F., BAEZ, J., 2005. "AGRE: Integrating Environments with Organizations." In D. WEYNS, V. D. PARUNAK, F. MICHEL, eds, *Environments for Multi-Agent Systems*. Berlin: Springer: 48–56.
- FISCHBACHER, U., GÄCHTER, S., 2010. "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *American Economic Review* 100(1): 541–56.
- FOUNTAIN, C., STOVEL, K., 2014. "Turbulent Careers: Social Networks, Employer Hiring Preferences, and Job Instability." In G. MANZO, ed., *Analytical Sociology: Actions and Networks*. Chichester: Wiley: 342–70.
- FRÍAS-MARTÍNEZ, E., WILLIAMSON, G., FRÍAS-MARTÍNEZ, V., 2011. "An Agent-Based Model of Epidemic Spread Using Human Mobility and Social Network Information." *Third International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, Boston 9–11 Oct., IEEE: 57–64.
Online: doi: 10.1109/PASSAT/SocialCom.2011.142.
- GALAN, J. M., IQUIERDO, L. R., 2005. "Appearances Can Be Deceiving: Lessons Learned Re-Implementing Axelrod's 'Evolutionary Approach to Norms'." *Journal of Artificial Societies and Social Simulation* 8(3): 2.
Online: <http://jasss.soc.surrey.ac.uk/8/3/2.html>.
- GALLEGATI, M., KIRMAN, A. P., 1999. *Beyond the Representative Agent*. Aldershot: Edward Elgar.
- GELMAN, A., HILL, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- GIGERENZER, G., GAISSMAIER, W., 2011. "Heuristic Decision Making." *Annual Review of Psychology* 62: 451–82.
- GILBERT, N., ABBOTT, A., 2005. "Introduction." *American Journal of Sociology* 110(4): 859–863.
- GILBERT, N., BANKES, S., 2002. "Platforms and Methods for Agent-Based Modeling." *Proceedings of the National Academy of Sciences* 99(supl. 3): 7197–8.
- GILBERT, N., TROITZSCH, K., 2005. *Simulation for the Social Scientist*, New York, NY: McGraw-Hill International.
- GINTIS, H., 2009. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton, NJ: Princeton University Press.
- GINTIS, H., 2013. "Markov Models of Social Dynamics: Theory and Applications." *ACM Transactions on Intelligent System and Technology* 4(3): 53.
Online: doi: 10.1145/2483669.2483686.

- GIRARDIN, L., CEDERMAN, L.-E., 2007. "A Roadmap to Realistic Computational Models of Civil Wars." In S. TAKAHASHI, D. SALLACH, J. ROUCHIER, eds, *Advancing Social Simulation: The First World Congress*. Berlin: Springer: 59–71.
- GOLDTHORPE, J., 2001. "Causation, Statistics, and Sociology." *European Sociological Review* 17(1): 1–20.
- GONZALEZ-BAILON, S., MURPHY, T. E., 2013. "Social Interactions and Long-Term Fertility Dynamics. A Simulation Experiment in the Context of the French Fertility Decline." *Population Studies* 67(2): 135–55.
- GOULD, R. V., 2002. "The Origins of Status Hierarchies: A Formal Theory and Empirical Test." *American Journal of Sociology* 107(5): 1143–78.
- GRANOVETTER, M., 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6): 1420–43.
- GRIMM, V., BERGER, U., DEANGELIS, D. L., POLHILL, J. G., GISKE, J., RAILSBACK, S. F., 2010. "The ODD Protocol: A Review and First Update." *Ecological Modelling* 221: 2760–8.
- GRÜNE-YANOFF, T., 2009. "The Explanatory Potential of Artificial Societies." *Synthese* 169(3): 539–55.
- HÄGERSTRAND, R., 1965. "A Montecarlo Approach to Diffusion." *Archives Européennes de Sociologie* 6(1): 43–67.
Online: doi: <http://dx.doi.org/10.1017/S0003975600001132>.
- HANSEN, L. P., HECKMAN, J. J., 1996. "The Empirical Foundations of Calibration." *Journal of Economic Perspectives* 10(1): 87–104.
- HEDSTRÖM, P., 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.
- HEDSTRÖM, P., BEARMAN, P., 2009. "What is Analytical Sociology all about?" In P. HEDSTRÖM, P. BEARMAN, eds, *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press: 3–24.
- HEGSELMANN, R., 2012. "Thomas C. Schelling and the Computer: Some Notes on Schelling's Essay on 'Letting a Computer Help with the Work'." *Journal of Artificial Societies and Social Simulation* 15(4): 9.
Online: <http://jasss.soc.surrey.ac.uk/15/4/9.html>.
- HELBING, D., 2012. *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior*. Berlin: Springer.
- HUET, S., DEFFUANT, G., 2008. "Differential Equation Models Derived from an Individual-Based Model Can Help to Understand Emergent Effects." *Journal of Artificial Societies and Social Simulation* 11(2): 10.
Online: <http://jasss.soc.surrey.ac.uk/11/2/10.html>.
- HUMMON, N. P., FARARO, T. J., 1995. "Actors and Networks as Objects." *Social Networks* 17(1): 1–26.
- IMHOFF, E. V., POST, W., 1998. "Microsimulation Methods for Population Projection." *Population* 10(1): 97–136.
- IZQUIERDO, A., IZQUIERDO, L. R. A., SEGISMUNDOS, A., 2013. "Combining Mathematical and Simulation Approaches to Understand the Dynamics of Computer Models." In E. BRUCE, M. Ruth, eds, *Simulating Social Complexity: A Handbook*. Berlin: Springer: 235–71.
- JACKSON, M., ed., 2012. *Determined to Succeed? Performance, Choice and Education*. Stanford, CA, Stanford University Press.
- JACKSON, M. O., 2008. *Social and Economic Networks*. Princeton, NJ: Princeton University Press.

- JANSSEN, M. A., ALESSA, L. N., BARTON, M., BERGIN, S., LEE A., 2008. "Towards a Community Framework for Agent-Based Modelling." *Journal of Artificial Societies and Social Simulation* 11(2): 6.
Online: <http://jasss.soc.surrey.ac.uk/11/2/6.html>.
- JONSSON, J. O., ERIKSON, R., 2000. "Understanding Educational Inequality: The Swedish Experience." *L'Année Sociologique* 50(2): 345–82.
- KIRMAN, A. P., 1992. "Whom or What Does the Representative Individual Represent?" *Journal of Economic Perspectives* 6(2): 117–136.
Online: doi: 10.1257/jep.6.2.117.
- KRONEBERG, C., KALTER, F., 2012. "Rational Choice Theory and Empirical Research. Methodological and Theoretical Contributions in Europe." *Annual Review of Sociology*, 38: 73–92.
- LAW, A., 2007. *Simulation Modeling and Analysis*. New York, NY: McGraw-Hill.
- LEOMBRUNI, R., RICHIARDI, M., 2005. "Why are Economists Skeptical about Agent-Based Simulations?" *Physica A* 355: 103–9.
- LYTINEN, S. L., RAILSBACK, S. F., 2012. "The Evolution of Agent-Based Simulation Platforms: A Review of NetLogo 5.0 and ReLogo." *Proceedings of the Fourth International Symposium on Agent-Based Modeling and Simulation*.
Online: http://condor.depaul.edu/slytinen/abm/Lytinen-Railsback-EMCSR_2012-02-17.pdf.
- MACAL, C. M., NORTH, M. J., 2013. "Successful Approaches for Teaching Agent-Based Simulation." *Journal of Simulation* 7(1): 1–11.
Online: doi:10.1057/jos.2012.1.
- MACHAMER, P. K., DARDEN, L., CRAVER, C. F., 2000. "Thinking about Mechanisms." *Philosophy of Science* 67(1): 1–25.
- MACY, M., FLACHE, A., 2009. "Social Dynamics from the Bottom Up: Agent-Based Models of Social Interaction." In P. HEDSTRÖM, P. BEARMAN, eds, *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press: 245–68.
- MACY, M., SATO, Y., 2008. "Reply to Will and Hegselmann." *Journal of Artificial Societies and Social Simulation* 11(4): 11.
Online: <http://jasss.soc.surrey.ac.uk/11/4/11.html>.
- MACY, M. W., WILLER, R., 2002. "From Factors to Actors: Computational Sociology and Agent-Based Modeling." *Annual Review of Sociology* 28: 143–66.
Online: doi: 10.1146/annurev.soc.28.110601.141117.
- MANZO, G., 2007. "Progrès et "urgence" de la modélisation en sociologie. Du concept de modèle générateur et de sa mise en œuvre." *L'Année Sociologique* 57(1): 13–61.
- MANZO, G., 2010. "Analytical Sociology and its Critics." *European Journal of Sociology* 51(1): 129–70.
- MANZO, G., 2013a. "Educational Choices and Social Interactions: A Formal Model and a Computational Test." *Comparative Social Research* 30: 47–100.
- MANZO, G., 2013b, *Educational Choices and Educational Traps. Towards an Integration between Computational and Statistical Modelling in the Sociology of Social Stratification*. FMSH-WP-2013-48, GeWoP1.
- MANZO, G., 2014. "Data, Generative Models, and Mechanisms: More on the Principles of Analytical Sociology." In G. MANZO, ed., *Analytical Sociology: Actions and Networks*. Chichester: Wiley: 4–52.
- MANZO, G., BALDASSARRI, D., 2015. "Heuristics, Interactions, and Status Hierarchies: An Agent-Based Model of Deference Exchange." *Sociological Methods & Research*, 44(2): 329–87.

- MARCHIONNI, C., YLIKOSKI, P., 2013. "Generative Explanation and Individualism in Agent-Based Simulation." *Philosophy of the Social Sciences* 43(3): 323–40.
- MATHIEU, P., BEAUFILS, B., BRANDOUY, O., eds, 2005. *Agent-Based Methods in Finance, Game Theory and their Applications*. Berlin: Springer.
- MEADOWS, M., CLIFF, D., 2012. "Reexamining the Relative Agreement Model of Opinion Dynamics." *Journal of Artificial Societies and Social Simulation* 15(4): 4. Online: <http://jasss.soc.surrey.ac.uk/15/4/4.html>.
- MILLER, J. H., PAGE, S. E., 2004. "The Standing Ovation Problem." *Complexity* 9(5): 8–16.
- MILLER, J. H., PAGE, S. E., 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.
- MONGE, P. R., CONTRACTOR, N., 2003. *Theories of Communication Networks*. New York, NY: Oxford University Press.
- MOSS, S., 2008. "Alternative Approaches to the Empirical Validation of Agent-Based Models." *Journal of Artificial Societies and Social Simulation* 11(1): 5. Online: <http://jasss.soc.surrey.ac.uk/11/1/5.html>.
- MOSS, S., EDMONDS, B., 2005. "Sociology and Simulation: Statistical and Qualitative Cross-Validation." *American Journal of Sociology* 110(4): 1095–131.
- MULDOON, R., 2007. "Robust Simulations." *Philosophy of Science* 54(5): 873–83.
- MÜLLER, B., BALBI, S., BUCHMANN, C. M. *et al.*, 2014. "Standardised and Transparent Model Descriptions for Agent-Based Models: Current Status and Prospects." *Environmental Modelling & Software* 55: 156–63. Online: [doi:10.1016/j.envsoft.2014.01.029](https://doi.org/10.1016/j.envsoft.2014.01.029).
- NAHIN, P. J., 2013. *Digital Dice: Computational Solutions to Practical Probability Problems*. Princeton, NJ: Princeton University Press.
- NIKOLAI, C., MADEY, G., 2009. "Tools of the Trade: A Survey of Various Agent Based Modeling Platforms." *Journal of Artificial Societies and Social Simulation* 12(2): 2. Online: <http://jasss.soc.surrey.ac.uk/12/2/2.html>.
- O'SULLIVAN, D., 2008. "Geographical Information Science: Agent-Based Models." *Progress in Human Geography* 32(2): 541–50.
- PADIOLEAU, J.-G., 1969. "Note sur les 'Simulations en sociologie politique'." *Revue Française de Sociologie* 10(2): 201–19.
- PAGE, L., GARBOUA, L. L., MONTMARQUETTE, C., 2007. "Aspiration Levels and Educational Choices: An Experimental Study." *Economics of Education Review* 26(6): 747–57. Online: [doi: 10.1016/j.econedurev.2007.06.001](https://doi.org/10.1016/j.econedurev.2007.06.001).
- POWERS, D. A., XIE, Y., 2000. *Statistical Methods for Categorical Data Analysis*. San Diego, CA/London: Academic Press.
- RAILSBACK, S. F., GRIMM, V., 2011. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton, NJ: Princeton University Press.
- RAUB, W., BUSKENS, V., VAN ASSEN, M. A. L. M., 2011. "Micro–macro Links and Microfoundations in Sociology." *Journal of Mathematical Sociology* 35(1–3): 1–25. Online: [doi:10.1080/0022250X.2010.532263](https://doi.org/10.1080/0022250X.2010.532263).
- REISS, J., 2011. "A Plea for (Good) Simulations: Nudging Economics toward an Experimental Science." *Simulation & Gaming* 42(2): 243–64.
- RICHIARDI, M., LEOMBRUNI, R., SAAM, N. J., SONNESSA, M., 2006. "A Common Protocol for Agent-Based Social Simulation." *Journal of Artificial Societies and Social Simulation* 9(1): 15. Online: <http://jasss.soc.surrey.ac.uk/9/1/15.html>.
- ROGERS, E. M., 2003. *Diffusion of Innovations*. New York, NY: Free Press.
- ROLFE, M., 2014. "Social Networks and Agent-Based Modelling." In G. MANZO, ed., *Analytical Sociology: Actions and Networks*. Chichester: Wiley: 237–70.

- ROSS, S. M., 2006. *Simulation*. Burlington, MA: Elsevier.
- ROUCHIER, J., 2013. "Construire la discipline "simulation agents". Les pistes proposées et un exemple." In F. VARENNE, M. SILBERSTEIN, eds, *Modéliser & simuler. Épistémologies et pratiques de la modélisation et de la simulation*, 2 vols. Paris: Éditions matériologiques: Ch. 29.
- SALTELLI, A. K., CHAN, E., SCOTT, M., 2000. *Sensitivity Analysis*. Chichester: Wiley.
- SAWYER, R. K., 2003. "Artificial Societies. Multiagent Systems and the Micro–Macro Link in Sociological Theory." *Sociological Methods & Research* 31(3): 325–63. Online: doi:10.1177/0049124102239079.
- SCHELLING, T. C., 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143–86.
- SHOHAM, Y., LEYTON-BROWN, K., 2009. *Multiagent Systems: Algorithmic, Game Theoretic, and Logical Foundations*. New York, NY: Cambridge University Press.
- SILVERMAN, E., BIJAK, J., HILTON, J., CAO, V. D., NOBLE, J., 2013. "When Demography Met Social Simulation: A Tale of Two Modelling Approaches." *Journal of Artificial Societies and Social Simulation* 16(4): 9. Online: <http://jasss.soc.surrey.ac.uk/16/4/9.html>.
- SMITH, E. R., CONREY, F. R., 2007. "Agent-Based Modeling: A New Approach for Theory Building in Social Psychology." *Personality and Social Psychology Review* 11(1): 87–104.
- SNIJDERS, T. A. B., 2011. "Networks Dynamics." In J. SCOTT, P. J. CARRINGTON, eds, *Sage Handbook of Social Network Analysis*. London: Sage: 501–13.
- SØRENSEN, A.B., 1976. "Models and Strategies in Research on Attainment and Opportunity." *Social Science Information* 15(1): 71–91.
- SQUAZZONI, F., 2012. *Agent-based Computational Sociology*. Chichester: Wiley.
- STOCKE, V., 2008. "Educational Decisions as Rational Choice? Testing the Erikson-Jonsson Model." *Working Papers 504*, Universität Mannheim.
- STONEDAHL, F., WILENSKY, U., 2010. "Evolutionary Robustness Checking in the Artificial Anasazi Model." In *Papers of the 2010 AAAI Fall Symposium on Complex Adaptive Systems: Resilience, Robustness, and Evolvability*, November 11–13, 2010, Arlington, VA.
- THIELE, J. C., KURTH, W., GRIMM, V., 2014. "Facilitating Parameter Estimation and Sensitivity Analysis of Agent-based Models: A Cookbook Using NetLogo and R." *Journal of Artificial Societies and Social Simulation* 17(3): 11. Online: <http://jasss.soc.surrey.ac.uk/17/3/11.html>.
- THORNE, B. C., BAILEY, A. M., PEIRCE, S. M., 2007. "Combining Experiments with Multi-Cell Agent-Based Modeling to Study Biological Tissue Patterning." *Briefings in Bioinformatics* 8(4): 245–57.
- THORNGATE, W., EDMONDS, B., 2013. "Measuring Simulation-Observation Fit: An Introduction to Ordinal Pattern Analysis." *Journal of Artificial Societies and Social Simulation* 16(2): 4. Online: <http://jasss.soc.surrey.ac.uk/16/2/4.html>.
- TJIMS, H., 2012. *Understanding Probability*. Cambridge: Cambridge University Press.
- TODD, P. M., BILLARI, F. C., SIMAO, J., 2005. "Aggregate Age-at-Marriage Patterns from Individual Mate-Search Heuristics." *Demography* 42(3): 5559–74.
- TRAIN, K. E., 2009. *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.
- TREUIL, J.-P., DROGOU, A., ZUCKER, J.-D., 2008. *Modélisation et simulation à base d'agents. Exemples commentés, outils informatiques et questions théoriques*. Paris: Dunod.

- TROITZSCH, K. G., 2014. "Analysing Simulation Results Statistically: Does Significance Matter?" In D. F. ADAMATTI, G. PEREIRA DIMURO, H. COELHO, eds, *Interdisciplinary Applications of Agent-Based Social Simulation and Modeling*. Hershey, PA: IGI Global: 88–105.
- VAN DE RIJT, A., SIEGEL, D., MACY, M., 2009. "Neighborhood Chance and Neighborhood Change: A Comment on Bruch and Mare." *American Journal of Sociology* 114(4): 1166–80.
- VARENNE, F., 2011. *Modéliser le social*. Paris: Dunod.
- VARENNE, F., 2013. "Modèles et simulations dans l'enquête scientifique: variétés traditionnelles et mutations contemporaines." In F. VARENNE, M. SILBERSTEIN, eds, *Modéliser & simuler. Épistémologies et pratiques de la modélisation et de la simulation*, 2 vols. Paris: Éditions Matériologiques.
- WANG, J., ZHANG, L., JING, C. *et al.*, 2013. "Multi-Scale Agent-Based Modeling on Melanoma and its Related Angiogenesis Analysis." *Theoretical Biology and Medical Modelling* 10: 41. Online: doi:10.1186/1742-4682-10-41.
- WILENSKY, U., RAND, W., 2007. "Making Models Match: Replicating an Agent-Based Model." *Journal of Artificial Societies and Social Simulation* 10(4): 2. Online: <http://jasss.soc.surrey.ac.uk/10/4/2.html>.
- WILL, O., HEGSELMANN, R., 2008. "A Replication That Failed—On the Computational Model in Michael W. Macy and Yoshimichi Sato: Trust, Cooperation and Market Formation in the U.S. and Japan. Proceedings of the National Academy of Sciences, 51 May 2002." *Journal of Artificial Societies and Social Simulation* 11(3): 3. Online: <http://jasss.soc.surrey.ac.uk/11/3/3.html>.
- WINSBERG, E., 2009. "A Tale of Two Methods." *Synthese* 169: 575–92.
- WINSHIP, C., MORGAN, S. L., 1999. "The Estimation of Causal Effects from Observational Data." *Annual Review of Sociology* 25: 659–706. Online: doi: 10.1146/annurev.soc.25.1.659.
- WIPPLER, R., LINDENBERG, S., 1987. "Collective Phenomena and Rational Choice." In J. ALEXANDER, B. GIESEN, R. MÜNCH, N. SMELSER, eds, *The Micro–Macro Link*. Berkeley, CA: University of California Press: 135–153.
- WOOLDRIDGE, M., 2000. *Reasoning about Rational Agents*, Cambridge, MA: MIT Press.
- WOOLDRIDGE, M., 2009. *An Introduction to Multi-Agent Systems*. Chichester: Wiley.
- YOUNG, H. P., 2006. "Social Dynamics: Theory and Applications." In K. JUDD, L. TESSFATSION, eds, *Handbook of Computational Economics*, vol. 2. Amsterdam: North Holland: 1081–108.
- ZHANG, L., WANG, Z., SAGOTSKY, J. A., DEISBOECK, T. S., 2009, "Multiscale Agent-Based Cancer Modeling." *Journal of Mathematical Biology* 58: 545–59.

To go further

At present, the sociologist eager to learn about agent-based simulation must train him- or herself by gathering different materials scattered within a heterogeneous literature. Indeed there is still no standard training course, codified in a manual for example, that delivers all the conceptual and technical expertise needed to work with this method. Similarly, training opportunities are rare. To overcome this limitation (accurately described by Charles M. Macal and Michael J. North, "Successful Approaches for Teaching Agent-Based Simulation," *Journal of Simulation*, 2013, 7: 1–11), I offer readers wishing to go further a series of stages amounting to a possible agent-based simulation self-study course.

Stage 1

Modelling and agent-based simulation are part of a research field, which is itself heterogeneous, that makes reference to "complexity." The first step is thus to familiarize yourself with the ideas and methods that shape this field. To this end, I suggest studying in the following order, Melanie Mitchell's general introduction (*Complexity. A Guided Tour*. Oxford: Oxford University Press, 2007) and the more specialist book by John H. Miller and Scott E. Page (*Complex Adaptive Systems. An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press, 2007).

Stage 2

To gradually work towards the more technical aspects, it is worthwhile first familiarizing yourself with the basic principles common to all computer simulation methods and second, to gain a better understanding of the way in which agent-based simulation differs from other forms of computer simulation. To do this I recommend reading, in the following order, Averill Law's book (*Simulation Modeling and Analysis*. New York, NY: McGraw-Hill, 2007) and Nigel Gilbert and Klaus Troitzsch (*Simulation for the Social Scientist*. New York, NY: McGraw-Hill International, 2005).

Stage 3

Armed with this more detailed, but still relatively general, knowledge of agent-based simulation, it is then worth familiarizing yourself with a few classic works in which this method was used to analyse theoretical models of specific social phenomena. So I advise you to study, in the following order, the review of models contained in Flaminio Squazzoni's book (*Agent-Based Computational Sociology*. Chichester: Wiley, 2012) as well as the many examples presented by Joshua M. Epstein (*Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press, 2006).

Stage 4

To complete your understanding of agent-based simulation, it is worth, finally, discovering its computational aspects. To do this I suggest making a detour through the literature on computing: the book by Michael Wooldridge (*An Introduction to Multi-Agent Systems*. Chichester: John Wiley & Sons, 2009) is an excellent starting point, as well as the very demanding book by Yoav Shoham and Kevin Leyton-Brown (*Multi-agent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge: Cambridge University Press, 2009).

.../...

Stage 5

An important element to be able to construct and study agent-based simulations is the modelling of social networks. So, before moving on to the more practical stage (the next stage), I suggest familiarizing yourself with the type of approach to social networks that best fits agent-based simulation. To this end, I recommend an introductory, but rigorous book by David Easley and Jon Kleinberg (*Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press, 2010).

Stage 6

At this stage, all the elements are in place for the reader to consider moving on to the practical work of agent-based simulation: programming. There are now many tools and languages for constructing multi-agent systems (for an overview, see Cynthia Nikolai and Gregory Madey, "Tools of the Trade: A Survey of Various Agent Based Modeling Platforms." *Journal of Artificial Societies and Social Simulation*, 2009, 12(2): 2, online: <http://jasss.soc.surrey.ac.uk/12/2/2.html>). Paradoxically, however, there are practically no books combining teaching a specific programming language, agent-based simulation and its application to social phenomena. Although flawed in several respects (in particular regarding algorithms enabling the construction of artificial networks) the book by Steven Railsback and Volker Grimm (*Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton, NJ: Princeton University Press, 2011) on the NetLogo language is a notable exception and is a very good starting point.

Stage 7

When an agent-based simulation has been built, it then has to be studied. A source providing a practical overview of the different strategies the researcher can pursue is the recent article by Jan C. Thiele, Winfried Kurth and Volker Grimm ("Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and R." *Journal of Artificial Societies and Social Simulation* 2014, 17(3): 11, online: <http://jasss.soc.surrey.ac.uk/17/3/11.html>). To understand the role of mathematical analysis in the study of agent-based simulation, the demanding book by Jean-Pierre Treuil, Alexis Drogoul and Jean-Daniel Zucker (*Modélisation et simulation à base d'agents. Exemples commentés, outils informatiques et questions théoriques*. Paris: Dunod, 2008) is a valuable reference.

Obviously, this training path is not the only possible one. It corresponds to my own way of learning, going from the general to the specific, as well as my own experience over the years. The proposed series of steps is, moreover, based on a necessarily reduced selection of resources currently available in the field of agent-based simulation. To expand on this selection and to complete my suggestions, I therefore invite the reader to explore the American economist, Leigh Tesfatsion's website at the following address: <http://www2.econ.iastate.edu/tesfatsi/ace.htm>.