European Journal of Science and Theology, September 2012, Vol.8, No.3, 103-113

EXPLAINING BY USING ARTIFICIAL SOCIETIES

Richard David-Rus*

Romanian Academy, Iasi Branch,Carol I bvd., no. 8, 700506, Iasi, Romania

(Received 4 April 2012)

Abstract

Computational models have an increasing impact in social and historical sciences. In this paper, I will focus on a specific type of modelling developed in computational social sciences, an agent-based model. My inquiry will aim to identify the sort of explanatory virtues that such a model could have. I will discuss the suggested possibility of causal explanations but also the recent proposal advanced by Grüne-Yanoff that sees them as potential functional explanations. In the last part I shall make some suggestions on how to advance our inquiry. These will point to the need of a more consistent explanatory context drawing on the way social scientists make use of such models. The other will point to the need of considering a cognitive process that is closely related to the explanatory one - understanding.

Keywords: simulation, explanation, models of historical evolution

1. A famous agent-based model used for the study of the historical evolution of a community

Agent-based models (ABM) simulate actions and interactions of autonomous agents (individuals or collective entities) with the goal to assess their effects on the system as a whole. They attempt to re-create and predict the appearance of complex phenomena and are characterized by emergence of new proprieties at macro level from microlevel dynamics. Early ideas can be found in Thomas Schelling's work [1] in the seventies and his classical now segregation model. But research explodes in the nineties with the increase in computational power. The book written by Epstein and Axtell, *Growing Artificial Societies* [2], is a landmark in the field and marks its maturation. They developed a large-scale ABM, the Sugarscape, to simulate and explore the role of the social phenomenon such as seasonal migrations, pollution, sexual reproduction, combat, and transmission of disease and even culture. The simulation that I will discuss further is a variant of this model.

The particular implementation of Sugarscape that I will use became one of the iconic models of the ABM community and attracted also significant attention in the media. It has as goal the simulation of a settlement history of Native Americans. The data refers to Ancestral Puebloans Indians also called Anasazi, who lived in North-east Arizona, in a delimited area called Long House Valley.

* E-mail: rusdavid@gmail.com

The specific geographic and climactic conditions made the valley an ideal place for preserving environmental records. The study simulates their settlements over a period of 500 years from 800 to 1350 AD. The initial model was developed by Dean and collaborators in 1999 [3] and was reproduced in later papers by other authors [4].

The agents in the model are individual households and their attribute are lifespan, movement capacities, nutrition requirements and storage capacities. The model takes as input data paleoenvironmental data: meteorological, groundwater and sediment deposition. From these data the production of maize per hectare for different categories of farming land. Secondly household movements (resettlement, land cultivation, procreation, annual deaths) are computed based on the previous results and the behaviour rules. The behavioural rules of the agents are provided by local ethnographical and anthropological studies of historic Pueblo people and other subsistence agriculturalists.

The output gives us the features of the history of the settlement, including population ebb and flow, changing spatial settlement patterns, and eventual decline. Simulated population levels closely follow the historical trajectory and reproduce also the dynamics of the clustering of the settlements (which tend to concentrate in the northern part of the valley). The original model [3] employs fairly homogenous agent attributes and reproduces the qualitative features of the history but it yields populations that were substantially too large. The follow-up paper [4] introduces heterogeneity for agent and landscape characteristics. This targets individual agents' onset of fertility, household fission and death, and harvest per hectare that are drawn from uniform distributions. The move improves the 'fit' of the model to the historical record. Nevertheless around year 1300 there is a remaining population in the simulation but not in reality. This indicates that some "additional push and pull factors" [4] are not accounted for in the actual simulations.

2. What sort of explanation?

One might rightly question himself about the sort of explanation such simulations might provide us. The fact that simulations provide explanations seems to be clear for scientists developing such models. For Dean and collaborators the close fit of the generated data to the observed data "indicates explanatory power" [3, p. 180]. Or, being more explicit about explanation, he states that "ultimately, 'to explain' the settlement and farming dynamics of Anasazi society in Long House Valley is to identify rules of agent behavior that account for those dynamics" [3, p. 201]. In a subsequent paper, Axtell and collaborators also claim more generally that "to *explain* an observed spatiotemporal history is to specify agents that generate – or grow – this history" [4] and that their account on the Anasazi society "goes a long way toward *explaining* this history" [4].

These statements are rather general and informal. From the point of view of the philosopher of science that can place the issue in the context of the heavily debated topic of scientific explanation the above claims call for more qualification. In the first steps I will try to address the issue by tackling some first sight solutions and presenting also a recent proposal due to Grüne-Yanoff [5].

Let's look for the beginning at the components of the explanatory relation, i.e. the *explanans* and the *explanandum*. For Epstein [6] in such simulations the macrostructure of interest constitutes the *explanandum*. In our case the *explanandum* is the dataset that records the settlement history. The closer the simulation generates this history, the better the explanation. The *explanans*, according to Epstein, the microspecification that generates the macrostructure is the explanation, meaning the *explanans*. The microspecification has to be sufficient to generate the macrostructure. In case of many microspecifications, empirical adequacy is sought as criterion of selection. So in our case the *explanans* is constituted by the ingredients of the simulations the agents and the behavioural rules according to which they act.

From the perspective of the explanations already discussed in the literature, the descriptions above might approximate the search for some known type of explanation. One of the first candidates to be taken into consideration is the causal explanation. In fact we can find scientists [7], as well as philosophers [8] considering causal explanations to be the final explanatory goal of such simulations.

The causal type of explanation is one of the most analyzed in the literature and recent accounts provide more sophisticated interpretations. Nevertheless the main idea can be stated clear enough: a causal explanation implies identifying the causes or tracking the causal influence that produced the fact or event to be explained. So the causes explain their effects. The specific accounts in the literature try to explicate causality in its explanatory dimension. For the purpose of our analysis we do not need some detailed articulation but can rely on some general accepted truths.

Grüne-Yanoff [5] argues that the chances to get real causal explanation are in this case small. Taking into account that we will need some causal regularities in the *explanans* that should be true or well-confirmed, we are confronted with a problematic situation. In our case the behavioural rules are the analogues of causal regularities but we cannot validate them independent of the simulation. They are not documented, since the population does not exist anymore. They were formulated by analogy with rules documented in other social groups. But here we have a problem with the transfer that is only approximate since there is a strong variation of such rules in small agricultural societies. Due to the fact that direct evidence is not available in order to validate the rules, Grüne-Yanoff [5] discusses other possibilities for validations such as the existence of a well-confirmed theory or the validation through behaviour experiments. He rejects them as incapable of establishing these rules.

The above lack of validation brings us to an alternative. A more plausible solution would rather take into consideration potential causal explanations. Unlike actual explanations, potential explanations lack validation of the law like generalizations [9]; they might be true but also they could be false. Another way to put it is to say that such simulations would provide possible causal histories, modalities in which the generated phenomena could have possibly been produced.

Grüne-Yanoff [5] dismisses this interpretation mainly on the following ground. The main problem would be that they are too numerous since we can obtain a great number of possible histories only by a simple variation of the parameters. One would have to select from these the genuine ones but criteria of selection are not available and the different accounts on causal explanation fail to provide any such criteria. The clarification of the term potential is therefore problematic and, besides Hempel's account, others accounts do not provide an explication of the idea.

Before proceeding to the discussion of another sort of solution, I want to address the above claims a little further**.** I think that they need to be nuanced. In regard to the previous argument against potential causal explanations one might argue that it assumes a sort of uniformity of the potential histories. Much of the huge number of solutions is generated by fine tuning of the parameters, which do not make a difference in the interpretation of the model. We might more pertinently talk of classes of such possibilities that are determined rather by ranges of variation than by the singular values. This would be the case for combinations too. So the filter that Grüne-Yanoff is asking for is given in the context of the specific inquiry in which the simulation is deployed.

Grüne-Yanoff [5] touches on this last idea when he discusses the possibility that the pragmatics of question can yield such criterion, as suggested by Lipton [10]. If the why-questions are contrastive, than the class of contrasts would induce the reduction of the set of possible histories, providing a refined set of possible histories. But Grüne-Yanoff [5] dismisses this by pointing to the practice of the computer scientists who are not asking such questions. This is true if we confine ourselves to the explanatory practice of computer scientists. Nevertheless if we are to consider the other scientists which are in fact rightly entitled to the explanatory business of social phenomena, we surely have to consider such questions that will induce the needed restrictions. It seems to me pertinent that the potential explanatory virtues of such simulations should be considered relative to the inquirers driven by explanatory goals in social phenomena, i.e. by social scientists.

Returning to the first argument against the possibility of a causal explanation, one might notice that it hinges on the validation of the behavioural rules. Nevertheless a recent replication of the simulation run by Janssen [11] that confirm the results of original paper revealed also some new aspect. According to it the specific agent behaviour contributes only a modest improvement to the sort of null model in which only the carrying capacity is considered. So whatever the rules would be from the range of variation that we can adopt, they do not have a major impact in the reproduction of the macropattern. The carrying capacity is defined as “the amount of households possible on the landscape based on the number of cells that produces enough food for one household” [11]. The models based on the carrying capacity is dependent only on two parameters the harvest variance and the harvest adjustment. These are the ones that have an impact on the *generandum* pattern.

What does this tell us in regard to the causal claims discussed above is that it makes even less sense to ask for causal detailed explanation in which the agent behaviour is involved. Rather on the basis of this model we can look for the sort of dependency between the major factor as the carrying capacity and the macropattern, which can be thought as a case of singular causality. This might undercut Grüne-Yanoff‘s critique that makes appeal to the causal regularities involved. As I will discuss in the last section we might see it rather as an explanatory suggestion that could be further worked out in the proper context of inquiry.

3. The functional interpretation

Before engaging into a further more detailed discussion of the possible explanatory virtues of this model, I will need to shortly present Grüne-Yanoff‘s own solution for the explanatory potential of this sort of simulation. He proposes an interpretation in terms of what in the literature on explanation is known as a functional explanation. Functional explanations were discussed from the very beginning as a particular type of explanation found in Biology and Social sciences. Hempel [9, p. 297-330] and Nagel [12] regarded such explanations as sketchy deductive-nomological explanations in the sense that when fully spelled out they instantiate the Hempelian schema. According to their analyses such explanations are meant to explain the existence of a component of a system by reference to the function it performs as a component of that system. To illustrate this we can mention the explanation of the existence of hearts in the organisms by indicating their functions in the organism, i.e.to pump blood through the circulatory system; or in Social sciences we might point to the social explanation that identifies the function of religious rituals to be the increase of the cohesion in society. The general form of a functional explanation can be rendered in statements as “The function of A in a system S with organization C is to enable S in environment E to engage in process P” and is reconstructed by Nagel as “every system S with organization C and in environment E engages in process **P;** if S with organization C and in environment E does not have **A,** then S does not engage in P; hence, S with organization C must have **A**” [12, p. 403].

This classical interpretation was disputed and alternatives were proposed. Grüne-Yanoff prefers another version developed by Cummins [13] in the seventies. In Cummins’ approach the performance of a system capacity is explained in terms of the capacities of the components it contains, and how they are organized. Such a functional analysis proceeds in two stages. In the first stage (the analytical stage) the function in question is analyzed in terms of the capacities involved in bringing about the function. In the second stage one shows that there is a physical structure present that realizes the various capacities. In order to illustrate this analysis we can make appeal to the explanation of how colour vision forms in humans. The process of formation of vision is decomposed into the various capacities of the anatomical components within the eye and nervous system that underlie the function (individual pigments in the cones, the molecular components of the photopigments and how they respond to light).

Two observations related to Cummins' account are in place here. First, in comparison to Nagel's account, the functional explanation in Cummins' analysis does not explain the existence of a component; it explains rather how a component contributes to the capacity of the system and in this sense exposes its causal contribution. The position was in fact described as a causal role account on functional explanation. And secondly, the larger capacity is not seen anymore as related to survival or reproduction or achieving some goal.

In applying this sort of analysis to Anasazi simulation we can proceed by identifying the elements of the analysis. So in our case the system's capacity to be analyzed is the capacity to generate the 'population dynamic' from the data set describing the meteorological and soil conditions. The component subsystems are: the households, settlement areas and farming plots, etc. Their capacities: movement, fertility, housing, crop yields, etc. The model organizes these capacities in a specific 'program', which is expressed through the behavioural rules of the households, the yield functions of the farming plots, etc., so that their combined operation, when fed with input data, produce the population data.

Nevertheless, in the end, due to the fact that the simulation does not replicate the extinction of the population from the valley after year 1300, Grüne-Yanoff considers that the model shortly fails to provide a potential functional explanation.

4. Some suggestions for a further working agenda

I will first discuss the Grüne-Yanoff's proposal and, as I unfold it, I will engage in some further investigations. This will ultimately bring me to some suggestions that I see as plausible for pursuing a viable solution.

A first concern comes from taking into consideration the previously mentioned result obtained by Janssen. The potential functional analysis that Grüne-Yanoff sees as being provided by such a simulations makes evident how different capacities of the components are related and contribute to the capacity of the system. Janssen's analysis showed that a single element, the carrying capacity, could account in a determinate way for the system's capacity. This reduces the explanatory consistency of the analysis considerably. Since all the other components might be eliminated and the emergent macropattern could be reproduced, we might have an almost spurious explanation by adding these components. They do not add much value to the functional explanation of the null model.

A second concern is related to the specificity of the system analyzed. The existence of different sorts of accounts in the literature raises the natural question regarding the possibility of subsuming them under one single concept. Philip Kitcher [14] proposed the concept of design as the one that can unify all the other analyses, but his position did not remain unchallenged. Godfrey-Smith [15] argued for a different perspective, a pluralistic one. For him there is no such unifying concept in all biological discourses in which functions are involved. Different areas of science require different analyses and therefore a better strategy is to adopt a pluralistic view. The different notions capture different kinds of information that scientists are looking for in different investigations. Godfrey-Smith position seems to me to be a more plausible one in the face of the variety of scientific practices and the complexity of the systems explored. In this sense, as Cummins' causal notion is more appropriate for such biological disciplines as Physiology where the scientists try to understand the dependency of the capacity of the system from the ones of its components. Meanwhile, in the Evolutionary biology or Behavioural ecology where the researchers' interests are directed towards the behaviour and structures of the organisms, the evolutionary notion is the appropriate one.

So, in order to determine the appropriateness of the used concept we might ask about the specificity of the domain in case of our system. What kind of system are we functionally analyzing? What kind of explanatory information are the scientists after when they use these models. At first sight it seems that it is an analysis of a social system since artificial societies are been generated. But one could probably talk of some ecological system as the target system of the computational model. On the other side, we are also dealing directly with a system that is constructed through a computer program, a simulation. We have also a functional analysis of the system constitute by the computer program. The functional analysis is taking place through the simulation. This stands in a representational relation to the real historical process, to the target system sharing the same organizational proprieties. And the main interest rests actually in the functional analysis of this target system.

We can better illuminate the above distinction if we make use of a schema that I have discussed in other texts [16]. The schema was proposed by Hartmann and Frigg at the conference on *Philosophical Perspectives on Scientific Understanding*, Amsterdam, 25-27 August 2005 and it is intended to be a rough account of the process of explanation through modelling. In their account they distinguish two different types of explanatory steps: one that takes place in the model and another that takes place when one transfers the knowledge obtained in the model to the real system. This can also be applied to a simulation which has in this case the same status as a model. The real explanatory load can rather be located in the second explanatory step. So the sciences are not interested *per se* in the derivation in the computer program in the simulation but in the transfer of this knowledge to the real system. The domain specificity is that of the target system and not of the computational system.

In the biological cases we could identify the sort of explanatory information the scientists were after, due to the fact that these episodes were embedded in a larger theoretical frame that determines the scientific problems investigate (physiological phenomena, evolutionary problems). It is not clear in our case what should such a frame be; the anthropological one being in this case a too general one. We will need a more delimited area of anthropological investigation with its specific problems. It is not also clear how we might transfer and claim appropriateness for a form of functional analysis from Biology to our system. A step forward in order to determine the specificity of the area would be to ask for more of an explanatory context and this implies placing such a model in a more specific scientific inquiry episode drawn from a scientific area that has properly explanatory goals. In such a context we could articulate the specific functional sort of analyses of the area. Or we can better argue for the plausibility of the transfer of the Cummins type of analysis and its appropriateness for the area investigated.

In this last part I'll advance some suggestions that depart from [which spring out of] the points discussed. The first one will make reference to a recent general account on explanatory power of simulation proposed by Weirich [17]. Under a more general view, which I think is shared by scientists too, we might identify the explanatory virtues of such a simulation in the fact that it points to factors at work in this situation. This is the idea that Weirich develops more carefully in his account. According to him, simulations provide explanation by identifying some of the factors at work in that situation. Rarely do they give us full explanations but usually they pick out only some of the factors and provide this way only partial explanation. But in order to be partial, an explanation should provide "an accurate account of some factors explaining the phenomenon; that is, it describes their interactions and effects, or their workings, with precision" [17, p. 159]. In this sense our case cannot be taken as a partial explanation since it lacks such precision and accuracy. Nevertheless, I think we can follow further the line and take into consideration also the possibility of such a partial explanation. So our case might rather be seen as a possible partial explanation, suggesting some possible factors at work in the situation. This will refer to both previous discussed sorts of explanatory claims, the potential causal one or the functional one.

But I would like to add to the above idea another twist. It draws on an aspect that had to be reflected more seriously in our considerations of explanatory virtues of simulations. This aspect considers the place of this practice in the larger context of inquiry. A proper approach on explanation should conceive it as embedded in the larger process of inquiry. In this sense, we could better view these practices as providing explanatory suggestions and not only possible explanations in the sense of explanations that will be or not validated. These suggestions represent modalities that can be followed in the further steps of investigation. In this way they could be seen as pointing to directions that could be further pursued in the investigation.

So such a simulation would make rather some explanatory suggestions. It indicates possible ingredients for an explanation and possible relations among them. It leaves open any further detailed articulation that has to be undertaken in a proper context. These factors could be further used in a functional explanation or in a causal mechanism explanation of a social system. The simulation does not impose a choice – so it remains an explanatory open suggestion that could be read in different keys. One of these keys is a causal one. Concentrating on a specific factor it might turn out in a proper investigation context that provides the additional information, to have a higher causal impact. This makes out for an inquiry that seeks to articulate a full-blooded causal explanation. Another key could be a functional reading. In a proper context the scientist might rather be interested in the system as a whole and the functions that some components are performing. To make these readings we will need to flash out these suggestions in a specific context of inquiry in a particular area of research. We need in this sense to provide the right pragmatics at work in that situation.

A last important suggestion that I want to advance makes appeal to an important process connected to the one of explanation. Understanding was seen as closely linked to explanation and in the most classical approaches in the literature was conceived as its goal. As I have discussed in other texts [18] the new tendencies in the Philosophy of science challenged this view and uncovered a variety of relations between the two processes. The new frame opens the possibility of approaching understanding in a direct way, without subsuming it to explanation. We could this way centre our inquiry on understanding directly and try to extract the possible explanatory consequences. The close relation between the two processes should suggest the explanatory potentials through the understanding gained in the modelling process.

Such a move facilitates also an approach that should look to the understanding that simulations provide us in the modelling context. If this understanding is of an explanatory sort, it should be qualified according to the benefits it brings to the scientific inquiry. In an attempt to decouple the issue of understanding from the one of explanation, Lipton [19] identifies more benefits that understanding could bring us without explanation. These are linked to causal information, that we can get through observation, experimentation, manipulation or inference, a sense of necessity provided through the fact that the process could not have been otherwise, a sense of what is possible that can come even from potential or false explanations and the unification obtained by comparing phenomena through analogies and classification. By considering understanding provided through simulation this way, we might actually identify modalities for a further search on full explanations. Overall, a strategy that brings in understanding as a closely related concept to explanation might prove to be more fruitful in the investigation of explanatory virtues of such computational models.

5. Conclusions

In order to draw my discussion to an end I will put the matters in a larger perspective and review the main points of my analysis. As we could see the issue of explanatory virtues of computational models is in general a well enough unexplored topic. Even in the frame of the recent philosophical research on models and simulations it only received some limited attention. This does not imply that it is a subject that could be ignored. On the contrary, in the face of the growing role such models play in the scientific practices and the explanatory intentions scientists associate with them, the philosophers have to consider it more seriously. In order to do this properly it is possible that we have to give up some of the classical assumptions on how to approach explanations and engage more exploratory strategies of inquiry.

Artificial societies occupy an important place in the computational social research. With reference to such a model I have tried to discuss some options that we might have when searching for the explanatory virtues of simulations. Opting for a potential sort of explanation seemed to be a plausible approach, through restricting it to only some sort of explanations, the functional one as Grüne-Yanoff does, didn't seem to me to be beyond any doubt. I have suggested another approach to the issue through essential two ways. First in order to raise properly the question concerning the explanatory virtues of such models, we would need a more scientifically consistent context, that could be embedded in an inquiry of a relevant scientific domain. This should belong to the scientific register specific to the nature of the phenomenon inquired, i.e. the social phenomenon. Secondly, I think that more of the explanatory value searched for could be exposed by making use of a related epistemic benefit – the understanding that scientists gain from such models.

In the end, any further inquiry into explanatory potential of simulations has to be done with reference to concrete examples from disciplines that target the specific phenomenon represented in the models and less to the formal analysis of such a system.

Acknowledgement

This paper is supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU ID 56815.

References

[1] T. Schelling, The Journal of Mathematical Sociology, **1(2)** (1971) 143.

[2] J. Epstein and R. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up*, Brookings Institution Press, Washinghton, 1996.

[3] J. Dean, G. Gumerman, J. Epstein, R. Axtell, A. Swedland, M. Parker and S. McCarrol, *Understanding Anasazi culture change through agent-based modeling*, in *Dynamics in human and primate societies: Agent based modeling of social and spatial processes*, T. Kohler and G. Gumerman (ed.), Oxford University Press, New York and Oxford, 1999, 179.

[4] R. Axtell, J. Epstein, J. Dean, G. Gumerman A. Swedland, J. Harburger, S. Chakravarty, R. Hammond, J. Parker and M. Parker, P. Natl. Acad. Sci. USA, **99** (2002) 7275.

[5] T. Grüne-Yanoff, Synthese, **169(3)** (2009) 539.

[6] J. Epstein, Complexity, **4(5)** (1999) 41.

[7] L. Tesfatsion, *Agent-based computational economics: A constructive approach to economic theory*, in *Handbook of computational economics*, vol. 2, L. Tesfatsion and K. Judd (ed.), Elsevier, North Holland, 2006, 831.

[8] K. Sawyer, Philosophical Explorations, **7(3)** (2004) 219.

[9] C. Hempel, *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, Free Press, New York, 1965, 338.

[10] P. Lipton, *Inference to the best explanation*, Routledge, London, 2001, 83.

[11] M. Janssen, Journal of Artificial Societies and Social Simulation, **12(4)** (2009) 13.

[12] E. Nagel, *The Structure of Science: Problems in the Logic of Scientific Explanation*, Harcourt, Brace & World, New York, 1961, 401.

[13] R. Cummins, *Journal of Philosophy*, **72** (1975) 741.

[14] P. Kitcher, *Midwest Studies in Philosophy*, **18(1)** (1993) 379.

[15] P. Godfrey-Smith, *Pacific Philosophical Quarterly*, **74** (1993) 196.

[16] R. David-Rus, *Logos & Episteme*, **2(2)** (2011) 177.

[17] P. Weirich, *Simulation & Gaming*, **42(2)** (2011) 155.

[18] R. David-Rus, *Explanation and Understanding through Scientific Models*, Doctoral thesis, Ludwig-Maximillians University, Munich, 2009, 15.

[19] P Lipton, *Understanding without Explanation*, in *Scientific understanding: philosophical perspectives*, H. de Regt, S. Leonelli and K. Eigner (eds.), University of Pittsburgh Press, Pittsburgh, 2009, 43.