

Different Modelling Purposes

Bruce Edmonds

Abstract:

How one builds, checks, validates and interprets a model depends on its ‘purpose’. This is true even if the same model is used for different purposes, which means that a model built for one purpose but now used for another may need to be re-checked, re-validated and maybe even re-built in a different way. Here we review some of the different purposes for building a simulation model of complex social phenomena, focussing on five in particular: theoretical exposition, prediction, explanation, description and illustration. The chapter looks at some of the implications in terms of the ways in which the intended purpose might fail. In particular, it looks at the ways that a confusion of modelling purposes can fatally weaken modelling projects, whilst giving a false sense of their quality. This analysis motivates some of the ways in which these ‘dangers’ might be avoided or mitigated.

Why read this chapter?

This chapter will help you understand the importance of clearly identifying one’s goal in developing and using a model, and the implications of this decision in terms of how the model is developed, checked, validated, interpreted and described. It might thus help you produce models that are more reliable for your intended purpose and increase the reliability of your modelling. It will help you avoid a situation where you partially justify your model with respect to different purposes but succeed at none of them.

Introduction

A common view of modelling is that one builds a ‘life-like’ reflection of some system, which then can be relied upon to act like that system. This is a correspondence view of modelling where the details in the model correspond in a one-one manner with those in the modelling target – as if the model were some kind of ‘picture’ of what it models. However, this view can be misleading since models always differ from what they model, so that they will capture some aspects of the target system but not others. With complex phenomena, especially social phenomena, it is inevitable that any model is, at best, a very partial picture of what it represents – in fact I suggest that this picture analogy is so unhelpful that it might be best to abandon it altogether as more misleading than helpful¹.

Rather, here I will suggest a more pragmatic approach, where models are viewed as *tools* designed and useful for specific purposes. Although a model designed for one purpose may turn out to be ok for another, it is more productive to use a tool designed for the job in hand. One may be able to use a kitchen knife for shaping wood, but it is much better to use a chisel. In particular, I argue that even when a model (or model component) turns out to be useful for more than one purpose it needs to be justified and judged with respect to *each* of the claimed purposes *separately* (and it will probably require re-coding). To extend the previous analogy: a tool with the blade of a chisel but the handle of a kitchen knife may satisfy some of the criteria for a tool to carve wood and some of the criteria for a tool to carve cooked meat, but fail at both. If one *did* come up with a new tool that is good at both, this would be because it could be justified for each purpose separately.

In his paper ‘Why Model?’, Epstein (2008) lists 17 different reasons² for making a model: from the abstract, ‘discover new questions’, to the practical ‘educate the general public’. This illustrates both the usefulness of modelling but also the potential for confusion. As Epstein points out, the power of modelling comes from making an informal set of ideas formal. That is,

¹ With the exception of the purpose of description where a model is *intended* to reflect what is observed.

² He discusses ‘prediction’ and then lists 16 other reasons to model.

they are made precise using unambiguous code or mathematical symbols. This lack of ambiguity has huge benefits for the process of science, since it allows researchers to share, critique, and improve models without transmission errors (Edmonds 2010). However, in many papers on modelling the purpose that its model was developed for, or more critically, the purpose under which it is being presented is often left implicit or confused. Maybe this is due to the prevalence of the ‘correspondence picture’ of modelling discussed above, maybe the authors conceive of their creations being useful in many different ways, or maybe they simply developed the model without a specific purpose in mind. However, regardless of the reason, the consequence is that readers do not know how to judge the model when presented. This has the result that models might avoid proper judgement – demonstrating partial success in different ways with respect to a number of purposes, but not adequacy against any.

Our use of language helps cement this confusion: we talk about a ‘predictive model’ as if it something in the code that makes it predictive (forgetting all the work in directing and justifying this power) – rather I am suggesting a shift from the code as a thing in itself, to code as a tool for a particular purpose. This marks a shift from programming, where the focus is on the nature and quality of the *code*, to modelling, where the focus is on the *relationship* of the behaviour of some code to what is being modelled. Using terms such as ‘explanatory model’ is OK, as long as we understand that this is shorthand for ‘a model which establishes an explanation’ etc.

Producing, checking and documenting code is labour intensive. As a result, we often wish to reuse some code produced for one purpose for another purpose. However, this often causes as much new work as it saves due to the effort required to justify code for a new purpose, and – if this is not done – the risk that time and energy of many researchers are wasted due to the confusions and false sense of reliability that can result. In practice, I have seen very little code that does not need to be re-written when one has a new purpose in mind. Ideas can be transferred and well-honed libraries for very well defined purposes, but not the core code that makes up a model of complex social phenomena³.

In this chapter, I will look at five common modelling purposes: prediction, explanation, theoretical exposition, description and illustration⁴. Each purpose is motivated, defined, and illustrated. For each purpose a ‘risk analysis’ is presented – some of the ways one might fail to achieve the stated purpose – along with some ways of mitigating these risks. In the penultimate section, some common confusions of purpose are illustrated and discussed, before the chapter concludes with a brief summary and plea to make one’s purpose clear.

Prediction

Motivation

If one can reliably predict anything that is not already known, this is undeniably useful regardless of the nature of the model (e.g. whether its processes are a reflection of what happens in the observed system or not⁵). For instance, the gas laws (stating e.g. that at a fixed pressure, the increase in volume of gas is proportional to the increase of temperature) were discovered long before the reason why they worked.

However, there is another reason that prediction is valued: it is considered the gold standard of science – the ability of a model or theory to predict is taken as the most reliable indicator of a model’s truth. This is done in two principle ways: (a) model A fits the evidence better than model B – a comparative approach⁶ or (b) model A is falsified (or not) by the evidence – a falsification approach. In either, the idea is that, given a sufficient supply of different models,

³ I am not ruling out the possibility of re-usable model components in the future using some clever protocol, it is just that I have not seen any good cases of code reuse, and many bad ones.

⁴ A later chapter (chap. ??) takes a more fine-grained approach in the context of understanding human societies

⁵ It would not really matter even if the code had a bug in it, if the code reliably predicts (though it might impact upon the knowledge of *when* we can rely upon it or not).

⁶ Where model B may be a random or null model but also might be a rival model.

better models will be gradually selected over time, either because the bad ones are discarded or outcompeted by better models.

Definition

By 'prediction', we mean *the ability to reliably anticipate data that is not currently known to a useful degree of accuracy via computations using the model.*

Unpacking this definition:

- It has to do it *reliably* – that is under some known (but not necessarily precise) conditions the model will work; otherwise one would not know when one could use it.
- The data it anticipates has to be *unknown* to the modeller. 'Predicting' out-of-sample data is not enough, since pressures to re-do a model and get a better fit are huge and negative results are difficult to publish.
- The anticipation has to be to a *useful* degree of accuracy. This will depend upon the purpose to which it is being put, e.g. as in weather forecasting.

Unfortunately, there are at least two different uses of the word 'predict'. Almost all scientific models 'predict' in the weak sense of being used to calculate some result given some settings or data, but this is different from correctly anticipating *unknown* data. For this reason, some use the term 'forecast' for anticipating unknown data and use the word 'prediction' for almost any calculation of one aspect from another using a model. However, this causes confusions in other ways so this does not necessarily make things clearer. *Firstly*, 'forecasting' implies that the unknown data is in the future (which is not always the case) and, *secondly*, large parts of science use the word 'prediction' for the process of anticipating unknown data. For example, if a modeller says their model 'predicts' something when they simply mean that it calculates it, then most of the audience may misunderstand and assume the author is claiming more utility than is intended.

As Watts (2014) points out, useful prediction does not have to be a 'point' prediction of a future event. For example, one might predict that some particular thing will not happen, the existence of something in the past (e.g. the existence of Pluto), something about the shape or direction of trends or distributions, or even qualitative facts. The important fact is that what is being predicted is not known beforehand by the modeller and that it can be unambiguously checked when it is known.

An Example

Nate Silver aims to predict social phenomena, such as the results of elections and the outcome of sports competitions. This is a data-hungry activity, which involves the long-term development of simulations that carefully see what can be inferred from the available data. As well as making predictions, his unit tries to establish the level of uncertainty in those predictions – being honest about the probability of those predictions coming about given the likely levels of error and bias in the data. As described in his book (Silver 2012) this involves a number of properties and activities, including:

- Repeated testing of the models against unknown data;
- Keeping the models fairly simple and transparent so one can understand clearly what they are doing (and what they do not cover);
- Encoding into the model aspects of the target phenomena that one is relatively certain about (such as the structure of the US presidential electoral college);
- Being heavily data-biased, requiring a lot of data to help eliminate sources of error and bias;
- Producing probabilistic predictions, giving a good idea about the level of uncertainty in any prediction;
- Being clear about what kind of factors are not covered in the model, so the predictions are relative to a clear set of declared assumptions and one knows the kind of circumstances in which one might be able to rely upon the predictions.

Post hoc analysis of predictions – explaining why it worked or not – is kept distinct from the predictive models themselves – this analysis may inform changes to the predictive model but is not then incorporated into the model. The analysis is thus kept independent of the predictive model so it can be an effective check. Making a good predictive model requires a lot of time getting it wrong with real, unknown data, and trying again before one approaches qualified successful predictions.

Risks

Prediction (as we define it) is very hard for any complex social system. For this reason, it is rarely attempted⁷. Many re-evaluations of econometric models against data that has emerged since publication have revealed a high rate of failure (e.g. Meese & Rogoff 1983) – 37 out of 40 model failed completely. Clearly, although presented as being predictive models, they did not actually predict unknown data. Many of these used the strategy of first dividing the data into in-sample and out-of-sample data, and then parameterising the model on the former and exhibiting the fit against the latter. Presumably, the apparent fit of the 37 models was not simply a matter of bad luck, but that all of these models had been (explicitly or implicitly) fitted to the out-of-sample data, because the out-of-sample data was known to the modeller before publication. That is, if the model failed to fit the out-of-sample data the first time the model was tested, it was then adjusted until it did work, or alternatively, only those models that fitted the out-of-sample data were published (a publishing bias). Thus, in these cases the models were not tested against predicting the out-of-sample data even though they were presented as such. Fitting known data is simply not a sufficient test for predictive ability.

There are many reasons why prediction of complex social systems fails, but two of the most prominent are: (1) it is unknown what processes are needed to be included in the model and (2) a lack of enough quality data of the right kinds. We will discuss each of these in turn.

1. In the physical sciences, there are often well-validated micro-level models (e.g. fluid dynamics in the case of weather forecasting) that tell us what processes are potentially relevant at a coarser level and which are not. In the social sciences this is not the case – we do not know what the essential processes are. Here, it is often the case that there are other processes that the authors have not considered that, if included, would completely change the results. This is due to two different causes: (a) we simply do not know much about how and why people behave in different circumstances and (b) different limitations of intended context will mean that different processes are relevant.
2. Unlike in the physical sciences, there has been a paucity of the kind of data we would need to check the predictive power of models. This paucity can be due to (a) there is not enough data (or data from enough independent instances) to enable the iterative checking and adapting of the models on new sets of unknown data each time we need to, or (b) the data is not of the right kind to do this. What can often happen is that one has partial sets of data that require some strong assumptions in order to compare against the predictions in question (e.g. the data might only be a proxy of what is being predicted, or you need assumptions in order to link sets of data). In the former case, (a), one simply has not enough to check the predictive power in multiple cases so one has to suspend judgement as to whether the model predicts in general, until the data is available. In the latter case, (b), the success at prediction is relative to the assumptions made to check the prediction.

A more subtle risk is that the conditions under which one can rely upon a model to predict well might not be clear. If this is the case then it is hard to rely upon the model for prediction in a new situation, since one does not know its conditions of application.

⁷ To be precise, some people have claimed to predict various social phenomena, but there are very few cases where the predictions are made public before the data is known and where the number of failed predictions can be checked. Correctly predicting events after they are known is much easier!

Mitigating Measures

To ensure that a model does indeed predict well, one can seek to ensure the following:

- That the model has been tested on several cases where it has successfully predicted data unknown to the modeller (at the time of prediction);
- That information about the following are included: exactly what aspects it predicts, guidelines on when the model can be used to predict and when not, some guidelines as to the degree or kind of accuracy it predicts with, any other caveats a user of the model should be aware of;
- That the model code is distributed so others can explore when and how well it predicts.

Explanation

Motivation

Often, especially with complex social phenomena, one is particularly interested in understanding *why* something occurs – in other words, explaining it. Even if one cannot predict something before it is known, you still might be able to explain it afterwards. This distinction mirrors that in the physical sciences where there are both *phenomenological* as well as *explanatory* laws (Cartwright 1983) – the former matches the data, the latter explains why that came about. In mature science, predictive and explanatory laws are linked in well-understood ways, but with less well understood phenomena one might have one without the other. For example, the gas laws that link measurements of temperature, pressure and volume were known before the explanation in terms of molecules of gas bouncing randomly around, and the formal connection between both accounts only made much later. Understanding is important for managing complex systems as well as understanding when predictive models might work. Whilst generally with complex social phenomena explanation is easier than prediction, sometimes prediction comes first (however if one can predict then this invites research to explain *why* the prediction works).

If one makes a simulation in which certain mechanisms or processes are built in, and the outcomes of the simulation match some (known) data, then this simulation can support an explanation of the data using the built-in mechanisms. The explanation itself is usually of a more general nature and the traces of the simulation runs are examples of that account. Simulations that involve complicated processes can thus support complex explanations – that are beyond natural language reasoning to follow. The simulations make the explanation explicit, even if we cannot fully comprehend its detail. The formal nature of the simulation makes it possible to test the conditions and cases under which the explanation works, and to better its assumptions.

Definition

By 'explanation' we mean *establishing a possible causal chain from a set-up to its consequences* in terms of the mechanisms in a simulation.

Unpacking some parts of this:

- The *possible causal chain* is a set of inferences or computations made as part of running the simulation – in simulations with random elements, each run will be slightly different. In this case, it is either a possibilistic explanation (A could cause B), in which case one just has to show one run exhibiting the complete chain, or a probabilistic explanation (A probably causes B, or A causes a distribution of outcomes around B) in which case one has to look at an assembly of runs, maybe summarising them using statistics or visual representations.
- For explanatory purposes, the structure of the model is important, because that limits what the explanation consists of. If, for example, the model consisted of mechanisms that are known *not* to occur, any explanation one established would be in terms of these non-

existent mechanisms – which is not very helpful. If one has parameterised the simulation on some in-sample data (found the values of the free parameters that made the simulation fit the in-sample data) then the explanation of the outcomes is also in terms of the in-sample data, mediated by these ‘magic’ free parameters⁸.

- The *consequences* of the simulations are generally measurements of the outcomes of the simulation. These are compared with the data to see if it ‘fits’. It is usual that only some of the aspects of the target data and the data the simulation produces are considered significant – other aspects might not be (e.g. might be artefacts of the randomness in the simulation or other factors extraneous to the explanation). The kind of fit between data and simulation outcomes needs to be assessed in a way that is appropriate to what aspects of the data are significant and which are not. For example, if it is the level of the outcome that is key then a distance or error measure between this and the target data might be appropriate, but if it is the shape or trend of the outcomes over time that is significant then other techniques will be more appropriate (e.g. Thorngate & Edmonds 2013)

Example

Stephen Lansing spent time in Bali as an anthropologist, researching how the Balinese coordinated their water usage (among other things). He and his collaborator, James Kramer, build a simulation to show how the Balinese system of temples acted to regulate water usage, through an elaborate system of agreements between farmers, enforced through the cultural and religious practices at those temples (Lansing & Kramer, 1993). Although their observations could cover many instances of localities using the same system of negotiation over water, they were necessarily limited to all their observations being within the same culture. Their simulation helped establish the nature and robustness of their explanation by exploring a close universe of ‘what if’ questions, which vividly showed the comparative advantages of the observed system that had developed over a considerable period. The model does not predict that such systems will develop in the same circumstances, but it substantially adds to the understanding of the observed case.

Risks

Clearly, there are several risks in the project of establishing a complex explanation using a simulation – what counts as a good explanation is not as clear-cut as what is a good prediction.

Firstly, the fit to the target data to be explained might be a very special case. For example, if many other parameters need to have very special values for the fit to occur, then the explanation is, at best, brittle and, at worst, an accident.

Secondly, the process that is unfolded in the simulation might be poorly understood so that the outcomes might depend upon some hidden assumption encapsulated in the code. In this case, the explanation is dependent upon this assumption holding, which is problematic if this assumption is very strong or unlikely.

Thirdly, there may be more than one explanation that fits the target data. So although the simulation establishes one explanation it does not guarantee that it is the only candidate for this.

Mitigating Measures

To improve the quality and reliability of the explanation being established:

- Ensure that the mechanisms built into the simulation are plausible, or at least relate to what is known about the target phenomena in a clear manner;
- Be clear about which aspects of the outcomes are considered significant in terms of comparison to the target data – i.e. exactly which aspects of that target data are being explained;

⁸ I am being a little disparaging here, it may be that these have a definite meaning in terms of relating different scales or some such, but too often they do not have any clear meaning but just help the model fit stuff.

- Probe the simulation to find out the conditions for the explanation holding using sensitivity analysis, addition of noise, multiple runs, changing processes not essential to the explanation to see if the results still hold, and documenting assumptions;
- Do experiments in the classic way, to check that the explanation does, in fact, hold for your simulation code – i.e. check your code and try to refute the explanation using carefully designed experiments with the model.

Theoretical exposition

Motivation

If one has a mathematical model, one can do analysis upon its mathematics to understand its general properties. This kind of analysis is both easier and harder with a simulation model – to find out the properties of simulation code one just has to run the code – but this just gives one possible outcome from one set of initial parameters. Thus, there is the problem that the runs one sees might not be representative of the behaviour in general. With complex systems, it is not easy to understand how the outcomes arise, even when one knows the full and correct specification of their processes, so simply knowing the code is not enough. Thus with highly complicated processes, where the human mind cannot keep track of the parts unaided, one has the problem of understanding how these processes unfold in general.

Where mathematical analysis is not possible, one has to explore the theoretical properties using simulation – this is the goal of this kind of model. Of course, with many kinds of simulation one wants to understand how its mechanisms work, but here this is the only goal. Thus, this purpose could be seen as more limited than the others, since some level of understanding the mechanisms is necessary for the other purposes (except maybe black-box predictive models). However, with this focus on just the mechanisms, there is an expectation that a more thorough exploration will be performed – *how* these mechanisms interact and *when* they produce different kinds of outcome.

Thus the purpose here is to give some more general idea of how a set of mechanisms work, so that modellers can understand them better when used in models for other purposes. If the mechanisms and exploration is limited this would greatly reduce the usefulness of doing this. General insights are what is wanted here.

In practice, this means a mixture of inspection of data coming from the simulation, experiments and maybe some inference upon or checking of the mechanisms. In scientific terms, one makes a hypothesis about the working of the simulation – why some kinds of outcome occur in a given range of conditions – and then tests that hypothesis using well-directed simulation experiments.

The complete set of simulation outcomes over all possible initialisations (including random seeds) does encode the complete behaviour of simulation, but that is too vast and detailed to be comprehensible. Thus some general truths covering the important aspects of the outcomes under a given range of conditions is necessary – the complete and certain generality established by mathematical analysis might be infeasible with many complex systems but we would like something that approximates this using simulation experiments.

Definition

‘Theoretical exposition’ means discovering then *establishing (or refuting) hypotheses* about the *general behaviour* of a set of mechanisms (using a simulation).

Unpacking some key aspects here.

- One may well spend some time illustrating the discovered hypothesis (especially if it is novel or surprising), followed by a sensitivity analysis, but the crucial part is showing these hypotheses are refuted or not by a sequence of simulation experiments.
- The hypotheses need to be (at least somewhat) general to be useful.

- A use of theoretical exposition can be to refute a hypothesis, by exhibiting a concrete counter-example, or to establish a hypothesis.
- Although any simulation has to have some meaning for it to be a model (otherwise it would just be some arbitrary code) this does not involve any other relationship with the observed world in terms of data or evidence.

Example

Schelling developed his famous model for a theoretical purpose. He was advising the Chicago district on what might be done about the high levels of segregation there. The assumption was that the sharp segregation observed must be a result of strong racial discrimination by its inhabitants. Schelling's model (Schelling 1969, 1971) showed that segregation could result from just weak preferences of inhabitants for their own kind – that even, a wish for 30% of people of the same trait living in the neighbourhood could result in segregation. This was not obvious without building a model, and Schelling did not rely on the results of his model alone but did extensive mathematical analysis to back up its conclusions.

What the model did not do is say anything about what actually caused the segregation in Chicago – it might well be the result of strong racial prejudice. The model did not predict anything about the level of segregation, nor did it explain it. All it did was provide a counter-example to the current theories as to the cause of the segregation, showing that this was not *necessarily* the case.

Risks

In theoretical exposition one is not relating simulations to the observed world, so it is fundamentally an easier and 'safer' activity⁹. Since a near complete understanding of the simulation behaviour is desired this activity is usually concerned with relatively simple models. However, there are still risks – it is still easy to fool oneself with one's own model. Thus the main risk is that there is a bug in the code, so that what one thinks one is establishing about a set of mechanisms is really about a different set of mechanisms (i.e. those including the bug).

A second area of risk lies in a potential lack of generality, or 'brittleness' of what is established. If the hypothesis is true, but only holds under very special circumstances, then this reduces the usefulness of the hypothesis in terms of understanding the simulation behaviour.

Lastly, there is the risk of over-interpreting the results in terms of saying anything about the observed world. The model might suggest a hypothesis about the observed world, but it does not provide any level of empirical support for this.

Mitigating Measures

The measures that should be taken for this purpose are quite general and maybe best understood by the community of simulators.

- One needs to check ones code thoroughly – see (Izquierdo et al 2017) for a review of techniques).
- One needs to be precise about the code and its documentation – the code should be made publically available
- Be clear as to the nature and scope of the hypotheses established.
- A very thorough sensitivity check, trying various versions with extra noise added etc.
- It is good practice to illustrate the simulation so that the readers understand its key behaviours but then follow this with a series of attempted refutations of the hypotheses about its behaviour to show its robustness.
- Be very careful about not claiming that this says anything about the observed world.

⁹ In the sense of not being vulnerable to being shown to be wrong later.

Description

Motivation

An important, but currently under-appreciated, activity in science is that of description. Charles Darwin spent a long time sketching and describing the finches he observed on his travels aboard the HMS Beagle. These descriptions and sketches were not measurements or recordings in any direct sense, since he was already selecting from what he perceived and only recording an abstraction of what he thought of as relevant. Later on, these were used to illustrate and establish his theoretical abstraction – his theory of evolution of species by natural selection.

One can describe things using natural language, or pictures, but these are inadequate for dynamic and complex phenomena, where the essence of what is being described is how several mechanisms might relate over time. An agent-based simulation framework allows for a direct representation (one agent for one actor) without theoretical restrictions. It allows for dynamic situations as well as complex sets of entities and interactions to be represented (as needed). This can make it an ideal complement to scenario development because it ensures consistency between all the elements and the outcomes. It is also a good base for future generalisations when the author can access a set of such descriptive simulations.

Definition

A description (using a simulation) is an attempt to *partially represent* what is important of *a specific observed case* (or small set of closely related cases).

Unpacking some of this:

- This is not an attempt to produce a 1-1 representation of what is being observed but only of the features thought to be relevant for the intended kind of study. It will leave out some features; in particular, it may leave out some of the interactions between processes.
- It is not in any sense general, but seeks to capture a restricted set of cases – it is specific to these and no kind of generality beyond these can be assumed.
- The simulation has to relate in an explicit and well-documented way to a set of evidence, experiences and data. This is the opposite of theoretical exposition and should have a direct and immediate connection with observation, data or experience.

Example

In (Moss 1998), Scott Moss describes a model that captures some of the interactions in a water pumping station during crises. This came about through extensive discussions with stakeholders within a UK water company about what happens in particular situations during such crises. The model sought to directly reflect this evidence within the dynamic form of a simulation, including cognitive agents who interact to resolve the crisis. This simulation captured aspects of the physical situation, but also tackled some of the cognitive and communicative aspects. To do this, he had represent the problem solving and learning of key actors, so he inevitably had to use some existing theories and structures – namely, Alan Newell and Herbert Simon’s “general problem solving architecture” (Newell & Simon 1972) and Cohen’s “endorsement mechanism” (Cohen 1984). However, this is all made admirably explicit in the paper. The paper is suitably cautious in terms of any conclusions, saying that the simulation “*indicate[s] a clear need for an investigation of appropriate organizational structures and procedures to deal with full-blown crises.*”

Risks

Any system for representation will have its own affordances – it will be able to capture some kinds of aspect much more easily than others will. This inevitably biases the representations produced, as those elements that are easy to represent are more likely to be captured than

those which are more difficult. Thus, the medium will influence what is captured and what is not.

Since agent-based simulation is not theoretically constrained¹⁰, there are a large number of ways in which any observed phenomena could be expressed in terms of simulation code. Thus, it is almost inevitable that any modeller will use some structures or mechanisms that they are familiar with in order to write the code. Such a simulation is, in effect, an *abduction* with respect to these underlying structures and mechanisms – the phenomena are seen through these and expressed using them.

Finally, a reader of the simulation may not understand the limitations of the simulation and make false assumptions as to its generality. In particular, the inference within the simulations may not include all the processes that are in what is observed – thus it cannot be relied upon to either predict outcomes or justify any specific explanation of those outcomes.

Mitigating Measures

As long as the limitations of the description (in terms of its selectivity, inference and biases) are made clear, there are relatively few risks here, since not much is being claimed. If it is going to be useful in the future as part of a (slightly abstracted) evidence base, then its limitations and biases do need to be explicit. The data, evidence or experience it is based upon also needs to be made clear. Thus, good documentation is the key here – one does not know how any particular description will be used in the future so the thoroughness of this is key to its future utility. Here it does not matter if the evidence is used to specify the simulation or to check it afterwards in terms of the outcomes, all that matters is that the way it relates to evidence is well documented. Standards for documentations (such as the ODD and its various extensions (Grimm et al. 2006, 2010) help ensure that all aspects are covered.

Illustration

Motivation

Sometimes one wants to make an idea clear, and an illustration is a good way of doing this. It makes a more abstract theory or explanation clear by exhibiting a concrete example that might be more readily comprehended. Complex systems, especially complex social phenomena, can be difficult to describe, including multiple independent and interacting mechanisms and entities. Here a well-crafted simulation can help people see these complex interactions at work and hence appreciate these complexities better. As with description, this purpose does not claim much, it is just a medium for the communication of an idea. If the theory is already instantiated as a simulation (e.g. for theoretical exposition or explanation) then the illustrative simulation might well be a simplified version of this.

Playing about with simulations in a creative but informal manner can be very useful in terms of informing the intuitions of a researcher. In a sense, the simulation has illustrated an idea to its creator. One might then exhibit a version of this simulation to help communicate this idea to others. However, this does not mean that the simulation achieves any of the other purposes described above, and it is thus doubtful whether that idea has been established to be of public value (justifying its communication in a publication) until this happens.

This is not to suggest that illustration is not an important process in science. Providing new ways of thinking about complex mechanisms or giving us new examples to consider is a very valuable activity. However, this does not imply its adequacy for any other purpose.

Definition

An illustration (using a simulation) is to communicate or make clear an idea, theory or explanation.

¹⁰ To be precise, it does assume there are discrete entities or objects and that there are processes within these that can be represented in terms of computations, but these are not very restrictive assumptions.

Unpacking this.

- Here the simulation does not have to fully express what it is illustrating, it is sufficient that it gives a simplified example. So it may not do more than partially capture the idea, theory or explanation that it illustrates, and it cannot be relied upon for the inference of outcomes from any initial conditions or set-up.
- The clarity of the illustration is of over-riding importance here, not its veracity or completeness.
- An illustration should not make any claims, even of being a description. If it is going to be claimed that it is useful as a theoretical exposition, explanation or other purpose then it should be justified using those criteria – that it seems clear to the modeller is not enough.

Example

In his book, Axelrod (1984) describes a formalised computational ‘game’ where different strategies are pitted against each other, playing the iterated prisoner’s dilemma. Some different scenarios are described, where it is shown how the “tit for tat” strategy can survive against many other mixes of strategies (static or evolving). The conclusions are supported by some simple mathematical considerations, but the model and its consequences were not explored in any widespread manner¹¹. In the book, the purpose of the model is to illustrate the ideas that the book proposes. The book claims the idea ‘explains’ many observed phenomena, but in an analogical manner – no precise relationship with any observed measurements is described. There is no validation of the model here or in the more academic paper that described these results (Axelrod & Hamilton 1981). In the academic paper there are some mathematical arguments which show the plausibility of the model but the paper, like the book, progresses by showing the idea is coherent with some reported phenomena – but it is the ideas rather than the model that are so related. Thus, in this case, the simulation model is an analogy to support the idea, which is related to evidence in a qualitative manner – the relationship of the model to evidence is indirect (Edmonds, 2001). Thus the role of the simulation model is that of an illustration of the key ideas and does not qualify for either explaining specific data, predicting anything unknown or exploring a theory.

Risks

The main risk here is that you might deceive people using the illustration into reading more into the simulation than is intended, as these are often quite persuasive in terms of their impact. Such simulations can be used as a kind of analogy – a way of thinking about other phenomena. However, just because you can think about some phenomena in a particular way does not make it true. The human mind is good at creating, ‘on the fly’, connections between an analogy and what it is considering – so good that it does it almost without us being aware of this process. The danger here is of confusing being able to think of some phenomena using an idea, and that idea having any force in terms of a possible explanation or method of prediction. The apparent generality of an analogy tends to dissipate when one tries to precisely specify the relationship of a model to observations, since an analogy has a different set of relationships for each situation it is applied to – it is a supremely flexible way of thinking. This flexibility means that it does not work well to support an explanation or predict well, since both of these necessitate an explicit and fixed relationship with observed data.

There is also a risk of confusion if it is not clear which aspects are important to the illustration and which are not. A simulation for illustration will show the intended behaviour, but (unlike when its theory is being explored) it has been tested only for a restricted range of possibilities, indeed the claimed results might be quite brittle to insignificant changes in assumption.

Mitigating Measures

Be very clear in the documentation that the purpose of the simulation is for illustration only, maybe giving pointers to fuller simulations that might be useful for other purposes. Also be

¹¹ Indeed the work spawned a whole industry of papers doing just such an exploration.

clear in precisely what idea is being communicated, and so which aspects of the simulation are relevant for this purpose.

Some Confusions of Purpose

It should be abundantly clear by now that establishing a simulation for one purpose does not justify it for another, and that any assumptions to the contrary risk confusion and unreliable science. However, the field has many examples of such confusions and conflation, so this message is obviously needed. It is true that a simulation model justified for one purpose might be used as *part* of the development of a simulation model for another purpose – this can be how science progresses. However, just because a model for one purpose *suggests* a model for another, does not mean it is a good model for the new purpose. If it is being suggested that a model can be used for a new purpose, it has to be justified for this new purpose. To drive home this point further, we look at some common confusions of purpose to underline this danger. Each time some code is mistakenly relied upon for a purpose other than has been established for it.

1. ***Theoretical exposition* → *Explanation***. Once one has immersed oneself in a model, there is a danger that the world looks like this model to its author. This is a strong kind of Kuhn's 'Theoretical Spectacles'¹², and results from the intimate relationship that simulation developers have with their model. Here the temptation is to jump from a theoretical exposition, which has no empirical basis, to an explanation of something in the world. A simulation can provide a way of looking at some phenomena, but just because one can view some phenomena in a particular way does not make it a good explanation. Of course, one can form a hypothesis from anywhere, including from a theoretical exposition, but it remains only a hypothesis until it is established as a good explanation as discussed above (which would almost certainly involve changing the model).
2. ***Description* → *Explanation***. In constructing a simulation for the purpose of describing a small set of observed cases, one has deliberately made many connections between aspects of the simulation and evidence of various kinds. Thus, one can be fairly certain that, at least, some of its aspects are realistic. Some of this fitting to evidence might be in the form of comparing the outcomes of the simulation to data, in which case it is tempting to suggest that the simulation supports an explanation of those outcomes. The trouble with this is twofold: (a) the work to test *which* aspects of that simulation are relevant to the aspects being explained has not been done; (b) the simulation has not been established against a range of cases – it is not general enough to make a good explanation. An explanation that only explains aspects of a small number of cases using a complex simulation is a bad explanation since there will be many other potentialities in the simulation that are not used for these few cases.
3. ***Explanation* → *Prediction***. A simulation that establishes an explanation traces a (complex) set of causal steps from the simulation set-up to outcomes that compare well with observed data. It is thus tempting to suggest that one can use this simulation to predict this observed data. However, the process of using a simulation to establish and understand an explanation inevitably involves iteration between the data being explained and the model specification – that is the model is fitted to that particular set of data. Model fitting is not a good way to construct a model useful for prediction, since it does not distinguish between what is essential for the prediction and the 'noise' (what cannot be predicted). Establishing that a simulation is good for prediction requires its testing against unknown data several times – this goes way beyond what is needed to establish a candidate explanation for some phenomena. This is especially true for social systems, where we often cannot predict events, but we can explain them after they have occurred.

¹² Kuhn (1962) pointed out the tendency of scientists to only see the evidence that is coherent with an existing theory – it is as if they have 'theoretical spectacles' that filter out other kinds of evidence.

4. **Illustration → Theoretical exposition.** A neat illustration of an idea suggests a mechanism. Thus, the temptation is to use a model designed as an illustration or playful exploration as being sufficient for the purpose of a theoretical exposition. A theoretical exposition involves the extensive testing of code to check the behaviour and the assumptions therein, an illustration, however suggestive, is not that rigorous. For example, it may be that an illustrated process is a very special case and only appears under very particular circumstances, or it may be that the outcomes were due to aspects of the simulation that were thought to be unimportant (such as the nature of a random number generator). The work to rule out these kinds of possibility is what differentiates using a simulation as an illustration from a theoretical exposition.

There is a natural progression in terms of purpose attempted as understanding develops: from illustration to description or theoretical exposition, from description to explanations and from explanations to prediction. However, each stage requires its own justification and probably a complete re-working of the simulation code for this new purpose. It is the lazy assumption that one purpose naturally follows from another that is the danger.

Conclusion

In Table 1 we summarise the most important points of the above discussion. This does not include all the risks of each kind of model, but simply picks the most pertinent ones.

Table 1. A brief summary of the discussed modelling purposes

Modelling Purpose	Essential features	Particular risks (apart from that of lacking the essential features)
<i>Prediction</i>	Anticipates <i>unknown</i> data	Conditions of application unclear
<i>Explanation</i>	Uses <i>plausible</i> mechanisms to match outcome data in a <i>well-defined</i> manner	Model is brittle, so minor changes in the set-up result in bad fit to explained data
<i>Theoretical exposition</i>	Systematically maps out or establishes the consequences of some mechanisms	Bugs in the code; inadequate coverage of possibilities
<i>Description</i>	Relates directly to evidence for a small set of cases	Unclear documentation; over generalisation from cases described
<i>Illustration</i>	Shows an idea clearly	Over interpretation to make theoretical or empirical claims

As should be clear from the above discussion, being clear about one's purpose in modelling is central to how one goes about developing, checking and presenting the results. Different modelling purposes imply different risks, and hence activities to avoid these. If one is intending the simulation to have a public function (in terms of application or publication), then one should not model with unspecified or conflated purposes¹³. Confused, conflated or unclear modelling purpose leads to unreliable models that are hard to check, can create deeply misleading results, and is hard for readers to judge – in short, it is a recipe for bad science.

¹³ This does not include private modeling, whose purpose maybe playful or exploratory, however in this case one should not present the results or model as if they have achieved anything more than illustration (to oneself). If one finds something of value in the exploration it should then be re-done properly for a particular purpose to be sure it is worth public attention.

Further Reading

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, 11(4), 12. <http://jasss.soc.surrey.ac.uk/11/4/12.html>

This gives a brief tour of some of the reasons to simulate other than that of prediction.

Edmonds, et al. (2017) *Understanding Human Societies*. This volume.

In this chapter, some modelling purposes that are specific to human social phenomena are examined in more detail giving examples from the literature.

Acknowledgements

I acknowledge you!

References

Axelrod, R. (1984), *The Evolution of Cooperation*, Basic Books,

Axelrod, R. & Hamilton, WD (1981) *The Evolution of Cooperation*. *Science*, 211:1390-1396.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford University Press.

Cohen, P. R. (1984). *Heuristic reasoning about uncertainty: an artificial intelligence approach*.

Cohen, P. R. (1984). *Heuristic reasoning about uncertainty: an artificial intelligence approach*. Pitman Publishing, Inc. Marshfield, MA.

Edmonds, B. (2001) *The Use of Models - making MABS actually work*. In. Moss, S. and Davidsson, P. (eds.), *Multi Agent Based Simulation, Lecture Notes in Artificial Intelligence*, 1979:15-32.

Edmonds, B. (2010) *Bootstrapping Knowledge About Social Phenomena Using Simulation Models*. *Journal of Artificial Societies and Social Simulation* 13(1)8. (<http://jasss.soc.surrey.ac.uk/13/1/8.html>)

Edmonds, B., Lucas, P., Rouchier, J. & Taylor, R. (2017) *Understanding Human Societies*. This Volume.

Epstein, J. M. (2008). Why model?. *Journal of Artificial Societies and Social Simulation*, 11(4), 12. <http://jasss.soc.surrey.ac.uk/11/4/12.html>

Grimm, V., et al. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*. 198:115–126.

Grimm, V., et al. (2010) *The ODD protocol: A review and first update*. *Ecological Modelling* 221:2760–2768.

Izquierdo, L. et al. (2017) *Checking Simulations*. This volume.

Kuhn, TS (1962) *The structure of scientific revolutions*, University of Chicago Press.

Lansing, JS & Kramer, JN (1993) *Emergent Properties of Balinese Water Temple Networks: Coadaptation on a Rugged Fitness Landscape*, *American Anthropologist*, (1):97-114.

Meese, RA & Rogoff, K (1983) *Empirical Exchange Rate models of the Seventies - do they fit out of sample?* *Journal of International Economics*, 14:3-24.

Moss, S. (1998) *Critical Incident Management: An Empirically Derived Computational Model*, *Journal of Artificial Societies and Social Simulation*, 1(4):1, <http://jasss.soc.surrey.ac.uk/1/4/1.html>

Newell, A., and Simon, H. A. (1972) Human problem solving Englewood Cliffs, NJ: Prentice-Hall.

Norling, E., Meyer, R & Edmonds, B. (2017) Informal Approaches to Developing Simulations. This volume.

Schelling, T. C. (1969). Models of segregation. *The American Economic Review*, 59(2), 488-493.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 1(2), 143-186.

Silver, N. (2012). *The signal and the noise: the art and science of prediction*. Penguin UK.

Thorngate, W. & Edmonds, B. (2013) Measuring simulation-observation fit: An introduction to ordinal pattern analysis. *Journal of Artificial Societies and Social Simulation*, 16(2):14 (<http://jasss.soc.surrey.ac.uk/16/2/4.html>).

Watts, D. J. (2014). Common Sense and Sociological Explanations. *American Journal of Sociology*, 120(2), 313-351.