Against Prior Theorising

Bruce Edmonds
Centre for Policy Modelling
Manchester Metropolitan University
Manchester, M1 3GH, UK.
http://www.cpm.mmu.ac.uk/~bruce

Experiment escorts us last -
His pungent company
Will not allow an Axiom
An Opportunity.
- *Emily Dickenson (1830–86)*

Introduction

Prior theory – that is theorising on the basis of thought and intuition[1], as opposed to attempting to explain observed data – inevitably distorts what comes after. It biases us in the selection of our data (the data model) and certainly biases any theorising that follows. It does this because we (as humans) can not help but see the world *through* our theorising – we are blind without the theoretical "spectacles" described by Kuhn (1962). If a theory has shown to be essentially correct in some domain (i.e. by thorough validation against the target problem or domain) using it as a framework can be helpful, however, if the theory is not mature or even speculative then it can effectively prevent progress[2]. I argue that, although we can not ever completely avoid this sort of bias, we can minimise its effect.

Two sources of prior theorising coming from opposite directions are sociology and formal systems – neither of these is *inherently* biased towards prior theorising, but just happens to be a source for such theorising at the present time. Computer scientists who *project* the results of interesting models *onto* society are also guilty of constructing first and fitting later.

Sociological Theory

Social simulators are sociologists that just happen to use modelling tools drawn from computer science and artificial intelligence, so the domain of study of sociologists and those in social simulation is the same. It thus is natural to look to sociology for theories to use in social simulation.

"Theory" has a different flavour in sociology than in many other sciences. It tends not to be a specifically applicable model but rather a rich sub-language of interrelated terms and ideas useful for characterising and understanding social phenomena. It certainly does not predict, and usually does not even explain, but it does frame and describe. This can be useful as a source of warnings about making unjustified assumptions, and can provide alternative ways of thinking about social phenomena. It also has a normative flavour, there being a strong expectation that one will put forward new data, theories, and models *within* the existing set of frameworks – Keith Sawyer is an example of this when he argues (in the e-mail debate) that social simulators *should* make use of sociology, but was unable to come up with any definite examples of social theory applicable to computational modelling or specific advantages in doing so.

However just because you *can* think about something in a particular way, or even that it *is* helpful to do so, does not mean it is actually *like that*, or that it is sensible to model it in that way. This is one of the reasons why sociology has been used so little in social simulation, *not* because the subject matter isn't relevant – it certainly is – but because of the lack of specificity to be found in most of its theory. It is not the macro/micro distinction, not the relevance, and not the domain that is the barrier but the *style* of current sociological theory. The trouble is that a linguistic framework, however rich, may not be

---

[1] A less polite term for prior theorising is "armchair theorising" (Simon 1986).

[2] For a discussion about the different characteristics of "revolutionary" science (before a well-validated theoretical framework appears) and "normal" science (after) see (Edmonds 2000b). For now all I will note is that in either case the *authority* of the science comes from the observation and not from the prior theory.

definite enough to allow a computational model to be constructed from it – more detail of the processes involved may be required for this. This is not completely for want of trying – for example, (Schillo et al. 2000) is a paper produced by a team that comprised of both computer scientists and sociologists, a paper which makes a brave attempt to bridge the gap between the thought of Bordieu and their specification for a social business agent.[3]

Sociology seems to have developed like this in order to escape the extremely specific contexts of its observations whilst at the same time rejecting the huge simplifying and reductionist assumptions necessary in order to apply traditional analytic modelling techniques. Both the semantic richness and the indefiniteness of sociological theory can be traced to this need for abstraction without reduction. What is needed for the development of *refutable theory* (or more precisely theory that can be effectively constrained by subsequent observation) is some *descriptive* modelling which marginally abstracts and generalises from the original observational contexts – I suggest that (the right kind of) computational modelling of a MABS type can precisely perform this role.

I contend that sociology *should* be producing theories that are amenable to computational animation in social simulations. This is not to say that other sociologists should do all the work for social simulators but that (more than a few) sociologists should be working towards this goal. That sociology almost completely fails to do this, is an damning indictment of its current state. The descriptive tools of social simulation may be able to play a part in developing sociology into a more productive discipline.

Formal Foundationalist Approaches

Many formal approaches to social phenomena are foundationalist in nature. That is they attempt to construct the formal *underpinnings* based upon a philosophical analysis of those phenomena. Sometimes this is accompanied by some suggestions as to how these formal constructions might be "brought to life" in a computational mechanism.

Many of these studies proceed on the basis of (some subset of) the following three arguments: *expressiveness*, *credibility* and *coherence*. The *expressiveness* argument goes roughly as follows: certain formal apparatus is necessary to cope with hypothetical situation *x* – if this is following other formal studies it is often in the form: existing suggestions of formal structures needs to be extended to be more expressive in order to cover situation *x*. Far too frequently *x* is a "toy" problem or example. *Credibility* is usually left implicit, it is that it is possible to draw an analogy between parts of the formal model and the target domain. For example in many studies which claim to use or address *emotion* one finds some reified object (such as a node) given the label of an emotion, and the reader is invited to interpret that object as such. *Coherence* is simply that it fits in with other, already established approaches – it is also usually left as implicit.

These three justifications are not enough to justify their inclusion in any particular domain of application, for there is no reason to suppose that they provide any "leverage" for the modeller or programmer. What these arguments *do* mean is that the problem domain can be expressed in the formal system (*expressiveness*) so that the reader can understand what they are supposed to mean (*credibility*) and can see how it fits in with other work (*coherence*).

Such a study hovers uncomfortably between pure mathematics and science. It would fail by the criteria applied in pure mathematics but it also does not justify its existence in the sense of aiding scientific effort. To qualify as a useful purely formal study it has to meet additional criteria, such as: *relevance*, *inferential leverage*, *tractability*, *generality*, and *explanatory power*. *Relevance* means that the system has been shown to be useful in modelling full-scale examples from the target domain (not just "toy" or "hypothetical" cases). *Inferential tractability* means that you can prove interesting theorems about the system (for example a translation into another system). *Tractability* means that the system is computationally tractable. *Generality* means that the formal system can be applied to a wide range of systems in the domain. *Explanatory power* means that the formal process in the system provide an explanation for some emergent process in the target domain. I discuss the role of formal systems and science in much greater depth in (Edmonds 2000b).

---

[3] I should make it clear that some of the case-studies found in sociology *are* extremely valuable and should be utilised to a far greater extent, but this useful source material is often buried in jargon. What would be even more useful is a set of ethnographic case-studies done on similar domains along *with* summarising statistical data to check computational models with. One problem is that one tends to get either good qualitative studies or statistical ones but not both!

Of course, many papers and research projects are not 100% foundationalist or representational but fall somewhere in between, for example (Conte and Paolucci 2001) is foundational in general approach but makes extensive use of real examples. The argument is still the same: that the research is more likely to be productive the more it is grounded in its domain of study and the less it is based on intuition.

Computer Scientists

There are still many computer scientists who build systems that are either loosely motivated by social phenomena or, in retrospect, can be interpreted that way, and then go on to interpret the behaviour of that system as if it told us something about how society actually works. In other words they project the behaviour of the system onto society in a way which is *far* stronger than is warranted. This is merely a less informed version of the sociological theory – the system provides a *way of thinking* about social phenomena and this is confused as information *about* the social phenomena. I discuss this problem in depth in (Edmonds 2000a), the paper which triggered this debate.

This type of computer science that is fundamentally divorced from real, messy, full-scale target domains and problems has been prevalent in the field since its inception. Perhaps computer science attracts people who prefer clean abstract worlds to the complex and contingent real one. There has always been a problem with computer programmers designing systems that are simply inadequate or irrelevant to their target domain. More recently the comparative[4] failure of GOFAI (Good Old-Fashioned Artificial Intelligence – a deliberately ironic acronym) once again illustrates how hubris (not accepting ideas from outside), prior theorising, and the lack of proper testing (prevalence of toy problems) is not a productive combination.

A Suggestion for a More Productive Way Forward

Rather that start with a general framework and progressively "fill in" the details, I am suggesting that it will be more profitable to start from the details and slowly abstract up from these. This is for several reasons: historically this sort of bottom-up approach is how major advances have been made (whilst successful top-down systematisation and formalisation has tended *follow* such advances); our domain of study is an extremely complex and contingent one so that premature guesses as to the shape of theory are less likely to be successful; and we are embedded within our subject matter so that our intuitions are more likely to be mislead.

Some of the desire to find short-cuts to "theory" before sufficient field-work is done, surely comes from the desire to emulate more established "hard" sciences. But whereas the popular impression is that these are *characterised* by their use of systematic formal structures, a *deeper* characterisation is the priority of data and its relationship with theory: data truly has the priority and can destroy theory that contradicts it (it acts to effectively constrain theory); a lot of effort is put into finding new ways of obtaining it with measurement and observational techniques; and, in practice (as opposed to how its often presented), the formal structures are used as no more than a flexible guide and tool.

Thus, I suggest, the following route to social theory will be more productive in the long run: start with many field-studies (observation); the development of ways to represent and summarise these in appropriate ways (generalised measurement and the formation of data models); the production of *descriptive* computational models based on these data models; and finally some abstraction of these descriptive models to more general constructions either computation or formal in kind.

Thus for social simulators the key is to produce descriptive models. General sociological theory and foundationalist formalism is entirely premature at this stage they will come later (if at all). What social simulators need is the field-studies and first-stage data models to base their models on and check them. Prior frameworks will more likely to mislead. Some examples of this sort of modelling are (Moss 1998, Hemelrijk 2000, Rouchier et al.).

Having to start from a pre-defined position (either "macro" or "micro") is a big disadvantage in this sort of enterprise – there have obviously co-evolved complex dependencies and causal loops between individuals and their society. This is one of the greatest opportunities for social simulation: models which have both levels so we can start to untangle the main interdependencies and interactions between

---

[4] Compared to the goals and promise of the fields that is – it is not obvious that it is *less* successful than many other fields, such as sociology.

the two! We can have it all. We do not have to decide beforehand and permanently which approach we are going to take, but can try to descriptively capture some cases where both are active, model them and see what contributes to constructing what phenomena, see what turns out to be the effective causal or explanatory principle in each case. The techniques of multi-agent based simulation allow this for the first time, before the only formal modelling techniques were necessarily summative. What beckons something beyond one position or another, something that mixes positions and gets closer to the mess that (individual and social) reality undoubtedly is.

It is possible that, in the future, some structuring and well-validated theoretical framework for broad classes of social phenomena may be discovered. In which case working within that framework will be a sensible approach[5]. However, this is far from the case at the moment and there seems little prospect of such a theory emerging in the near future. This is unsurprising – our target phenomena (the social world) is at least as complex as the biological world, less easy to study and highly context-dependent. It is likely that we will have to do much more empirical descriptive work before our Darwin or Newton appears.

A Concluding Illustration

To illustrate the various positions in this debate, let me transpose some the arguments to a more established domain and see how they appear. I chose biology because it is closest to the social domain in terms of complexity and contingency, but is established and successful. This is slightly unfair, but the point is to emphasise the *priority* of observation and data over theory and the *priority* of theory based on observation over other kinds such as simplified, systematised or merely attractive accounts. If we let ourselves follow this priority our theorising will be lead more by the target phenomena than our previous theorising, and so be more likely to lead to substantial progress.

A lot of sociology may be likened to Aristotelian Biology – a (now outmoded) philosophical framework for considering the biological world that prevailed up to the 18th century. Aristotelian Biology is credible, rich in meaning, and coheres well with other philosophy. It is designed on an essentially *a prior* basis – loosely linked to intuition but not originated in or based on any extensive field-work. Its practitioners exhorted people to learn it and relate any biological observation done *using* its framework. However it turned out not to further the field of biology, but rather impede its progress. It contrasts markedly with Darwinian Biology, which was born of painstaking observation. The establishment of Darwinism was messy, and it was only much later with the Darwinian/Genetic synthesis in the second half of the 20th century that a systematised account was possible. Just because we dismiss biology of the *a priori* sort does not mean that we need to dismiss empirically lead biology, just that it is not sensible to try and determine a complex and contingent domain by discursive thought only dimly inspired by observation and experience.

On the other hand, imagine trying to produce a formal axiomatic account of some aspect of biology, say the dermatology of rats, before extensive field-work has been done (in other words attempt a foundationalist study). One may indeed build up an elegant and believable theory based on the need of rats for protection from microbes, to keep in heat etc. One may even build simulations based on these theories (and get away with it as long as one only uses "toy" problems and domains to test them). One may build up a tradition where this theory is refined and extended. Yet it still need not bear *any* significant resemblance at all to the skins of rats as they happen to have evolved. Compare this to the situation once the huge amount of biological observation and experiment has been done, and a picture is built up of how rats skins *actually* are. First there is a lot of simple descriptive work, cross-sections, observation of rats in the wild and in captivity, experiments etc. and *then* this is abstracted into an account which delineates some of the functions of the skin as meaningful sub-systems (pain detection, sweat, heat control, hair production etc.). Now a somewhat simplified and systematised account is possible, separating out what is universal to rat's skins and what is contingent to particular rats, as well as what is relevant for certain purposes and what is not.

Finally, consider the argument that *societies only emerge from the interactions of individuals, and to specify society level phenomena will merely constrain our models of society in a reductionist way* (which is my summary of one of Rosario Conte's themes) transposed to biology. It becomes: *organisms only emerge from the interactions of cells, and to specify organism level phenomena (such as organs) will merely constrain our models of organisms in a reductionist way*. Clearly, despite the

---

[5] As long as its domain of applicability is kept in mind.

fact that we know that organisms started as single-celled creatures and only built up more complex *soma* over the millions of years of evolution, it would be daft to limit the modelling of organism-level phenomena to only those that were emergent from the interactions of cells. It is far better to: specify what one (thinks one) knows; guess at those aspects that are not obvious; and check the model output against what one observes. Far better, what ever the "level" these aspects happen to be, because we *know* that these aspects co-evolved with the rest of the organism's active functions to form meaningful and identifiable sub-systems. The pragmatics of modelling are more important than one's philosophical preference (Edmonds 1999).

To take positive lessons from biology. Biologists spend a lot of time doing field-work and little time doing formal modelling. The subject matter is not composed of neat systems but thousands of species, habitats, processes, and systems that have to be (to a large extent) categorised and modelled differently. Description prevails over theory. What theory there is came after a lot of descriptive work, and when it came was messy[6] – truly commensurate with its subject matter.

References

Edmonds, B. (1999). *Pragmatic Holism*, Foundations of Science, **4**:57-82.

Edmonds, B. (2000a). 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. Also available at: http://www.cpm.mmu.ac.uk/cpmrep74.html

Edmonds, B. (2000b) The Purpose and Place of Formal Systems in the Development of Science, CPM Report 00-75, MMU. Available at: http://www.cpm.mmu.ac.uk/cpmrep75.html

Conte, R. and Paolucci, M. (2001). Intelligent Social Learning. *Journal of Artificial Societies and Social Simulation*, **4**(1). http://www.soc.surrey.ac.uk/JASSS/4/1/3.html

Hemelrijk, C. K. (2000) Sexual Attraction and Inter-sexual Dominance among Virtual Agents, In. Moss, S. and Davidsson, P. (eds.), Multi Agent Based Simulation, *Lecture Notes in Artificial Intelligence*, **1979**:167-180.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

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

Rouchier J., Bousquet F., Requier-Desjardins M., and Antona M., (2001), A Multi-Agent Model for Describing Transhumance in North Cameroon: Comparison of Different Rationality to Develop a Routine, *Journal of Economic Dynamics and Control*, **25**:527-559.

Schillo, M., Fischer, K., Allen, S. and Klein, C. T. (2000) Socially Competent Business Agents with Attitude - Using Habitus Field theory to Design Agents with Social Competence. "Starting from Society" symposium, ASIB'2000, Birmingham, April 2000. http://www.cpm.mmu.ac.uk/cpmrep68.html

Simon, H.A. (1986). The failure of armchair economics [Interview]. *Challenge*, 29(5), 18-25.

[6] The theory of evolution (as presently understood) is by far from clean and straightforward.