How Formal Logic Can Fail to be Useful for Modelling or Designing MAS

Bruce Edmonds

Centre for Policy Modelling
Manchester Metropolitan University
`http://cfpm.org/~bruce`

"To a person with a hammer, every screw looks like a nail" (trad.)

Abstract. There is a certain style of paper which has become traditional in MAS – one where a formal logic is introduced to express some ideas, or where a logic is extended on the basis that it then covers certain particular cases, but where the logic is not actually *used* to make any substantial inferences and no application of the logic demonstrated. I argue that although these papers do follow a certain tradition, that they are not useful given the state of MAS and should, in future, be rejected as premature (just as if one had simulation but never run it). I counter the argument that theory is necessary by denying that the theory has to be so abstract. I counter the argument that logic helps communication on the simple grounds that for most people it doesn't. I argue that the type of logic that tends to be used in these papers is inappropriate. I finish with some suggestions as to useful ways forward.

Introduction

During RASTA 2002 there was some discussion about the utility of formal systems for building or understanding multi-agent systems (MAS). This paper is an attempt to put my arguments. I argue that (as with any tool) one has to use formal systems appropriately. Merely following a tradition of how to use and develop a particular kind of formal system is not sufficient to ensure one is doing something useful.

In this context I wish to make it clear that I have nothing against logic. I like formal logics because they can deal with qualitative information and they can be quite expressive. However, at the end of the day[1], they are just one of a range of types formal systems that could be used – the kind of the system that is chosen is important. The point is to distinguish when and how a particular formal system is useful – this applies to formal logics as a particular case.

In short, the question is not whether to abstract from our field of study using formal systems but how. In the past, premature 'armchair theorising' has not helped the eventual emergence of useful theory, but rather impeded it. Formal systems (such as logics) are not the *content* of theory but merely a *tool* for expressing and applying

[1] As David Hales would say.

theory in a symbolic way – choosing the wrong kind of formal system will bias our attempts and make our task more difficult.

Two 20th Century Trends in Logic

Whitehead and Russell (1962) showed that set theory, arithmetic and a good chunk of other mathematics could be formalised using first-order classical predicate logic. This dramatically demonstrated the expressive power of logic. Once set theory was properly logically formalised and the expressive power of set theory revealed it became clear that all mathematics could be embedded in set theory and hence be logically formalised. If any system could be shown to have an embedding in set theory, then it counted as mathematics. Thus set theory and classical first order predicate calculus was shown to *general systems*, in the sense that all known formal systems could be expressed in them (albeit with different degrees of difficulty).

In the second half of the 20th Century there was an explosion of different kinds of logic. This can be divided up into two approaches: those who were searching for the 'one true logic' (what I call the 'philosophical approach'); and those who saw logic as merely a useful tool for doing complex inference (what I call the 'pragmatic approach'). The former of these tinkered with the very structure of logic, restructuring the nature of deduction in the logic so as to match correct inference in natural language and by inventing new objects into the logic such as indices, operators, names etc. The nature of their discussions went very much by example – since they felt it was worth trying to construct the 'one true logic' it necessarily had to include all such cases. Logics in this vein included intuitionistic logic, free logic, relevance logic and modal logic. Due to the nature of their discussions their work tended to concentrate upon the axioms of the logic in relation to particular cases and treat the proof theory and formal semantics more as an after thought.

The pragmatic approach does not care so much about the philosophical interpretation as to what could be done with the logic. Thus, since classical first order predicate logic was generally expressive (Gabbay 1993), they tended to work within this framework or construct simple extensions of it. For these people it was the pragmatic virtues that mattered: was it good for doing inference in; were its formal semantics checkable; was it easy to model with; and could it be used for computation (ala Prolog and its successors)? The particular logic chosen for the MAS modelling language, SDML[2] is a case in point – its purpose is not to capture any general theory of cognition but to provide a sound and efficient basic for the consistent firing of complex sets of interdependent rules (Moss et al. 1998).

Unfortunately the philosophical approach has tended to attract the more attention in AI. There may be many reasons for this: it may be that the association with philosophy gives it academic status; it may be that the participants truly believe that there will be general logical systems that encode cognitive relations in ultimately simple ways; and it may be that it is relatively easy to do but difficult to criticise. Whatever the reason there has grown up a tradition in AI (and now MAS) which

[2] `http://sdml.cfpm.org`

discusses and compares different axiomatisations of logic and logical systems based purely on plausibility and the ability to encode particular examples (i.e. its expressive power). It is this approach that I am arguing against on the grounds that it will not be useful in either understanding or building MAS.

Generality and Abstraction

One of the principle ways of achieving generality is to abstract away from the detail of particular cases leaving only what happens to be true of the wider domain one is considering (post hoc abstraction). Another way is to decide the structure before hand and to *choose* one's domain accordingly or else to simply ignore those aspects of those cases that seem to contradict that structure (a priori abstraction). A third way is to include a method for adapting to the particularities of each case so that the detail is preserved (adaptive generality). However it is achieved, the increased generality is obtained at a cost, a cost of lost information, relevance or computation respectively. The cost of losing information as a result of post hoc abstraction may be critical if it is the important details (w.r.t. one's goal) that are lost. The cost of restricted relevance as a result of a priori abstraction may be critical if this means that it excludes your intended object of study. The cost of increased computation may be critical if the computation is too onerous to be practical.

One well-known dynamic of philosophical discourse is that of the counter-example followed by an increase in generality: a thesis is proposed; then a case exhibited where the thesis fails; and, in response, the thesis is generalised (e.g. by adding caveats, or by being suitably elaborated). The repeated application of this process of a priori abstraction is a set of very general, but irrelevant principles. These principles may give one the illusion of relevance because the 'ghosts' of the original concepts are left as labels and symbols in the general principles and one has the impression that the relevance can be restored by the simple adding of particulars. However, if this attempted this is found to be unworkable in practice. Be clear – it is not generality or abstraction by themselves that causes this lack of relevance but the *way* the generality is achieved (i.e. a priori abstraction). Similarly – I am not arguing against generality or abstraction but that it should be done in a way that results in useful theory. Work which attempts to mimic the counter-example-generalisation process in formal logic will not result in relevant theory about MAS.

One way of clearly demonstrating that increased generality is not a sufficient reason for exhibiting a logic is that there are already many logics (and other formal systems) that are as general as possible. If a particular logic has the ability to capture a particular concept then the general one will also be able to do this. The point of inventing new formal systems is thus *entirely* pragmatic, for each system (even the general ones) will inevitably facilitate the construction of certain systems and frustrate others, just as different programming languages are good at certain tasks and bad at others. This presence of implicit bias is not a question of the theoretical ability of the system but practical ease for us humans. This is why we neither formalise everything in set theory nor program using Turing Machines. Choosing an inappropriate formal system will bias the development of a theory in unhelpful ways, choosing an appropriate system will facilitate it (Edmonds 2000). Merely establishing that a particular system can express certain properties does not demonstrate that the system will facilitate a good theory, for the general systems also do this and they would (almost certainly) make formal modelling impossibly cumbersome and inference infeasible.

Thus arguing for a particular kind of formal logic on the grounds that it is able to express certain ideas, concepts or cases is very weak, for there are already formal logics that do the same (if any can). Thus, although the development of formal logics is often driven by a wish to express certain ideas, they need to be *justified* on other, stronger grounds.

The Need for Theory

Clearly if we are to escape simply considering individual cases and if our understanding of MAS is to inform our construction of MAS (and *vice versa*) then we will need to generalise and abstract our knowledge, i.e. use 'theory'. The trouble is that 'theory' can come in a variety of levels of abstraction and a variety of forms. A natural language description is already a sort of theory because it is the result of many relevance and representational decisions – it provides a level of generalisation by facilitating the comparison of phenomena by substituting the comparison of descriptions. An MAS may be also be used as a method of producing a sort of dynamic description of a social system – this is when one attempts to program the individual agents as closely to actual accounts as possible and then check that all stages also correspond to those in the social systems at all levels of aggregation. Another MAS may be intended to represent a set of phenomena that occurs in a small set of individual cases – here the generality is restricted to a particular domain. At the other end of the scale are the 'high theories' of philosophy or sociology – these are ideas that are supposed to have a very great level of generality. In philosophy the theories tend to be precise but irrelevant. In contrast, in sociology the theories are relevant but often extremely difficult to pin down – they are more akin to a richly expressive language for talking and thinking about social phenomena.

I am unsure of exactly what Rosaria Conte means by 'theory' during her remarks during the closing panel of AAMAS 2002 (and elsewhere). If she meant that *some* level of abstraction will be necessary for escaping from individual cases, then I agree with her – simply constructing particular MAS is not enough. However, if she is arguing that 'high theory' is necessary, then I disagree, for intermediate levels of abstraction also allows us to escape from single cases. For example physicists managed perfectly well to develop useful theories before the advent of their high theories, indeed they are still looking for a 'Theory of Everything' (TOE), even though it is clear that the situations in which such a TOE would diverge from the more mundane theories we already have will be extreme and unusual (from our point of view).

In the past theory that is mainly based on intuition which overtakes its evidential warrant has not had a good track record in resulting in useful theory. In fact, there is evidence that it has actively hindered the development of useful theory. A classic example of this is the thought of Aristotle on anatomy, which was wrong but played a part in delaying the spread of accurate information derived from dissection. Part of the reason for this is that theories play an important role in providing a language for thought, which (amongst other things) effects what evidence we look for (Kuhn 1962) and biases further modelling effects (since other kinds of models will probably not fit well in that framework).

Thus papers proposing 'high theories' of MAS need substantial justification before being trusted and certainly more than a few cases and vague intuitions. Further, such high theory is unnecessary in order to escape particular cases and experiences – models that are specific to particular kinds of MAS and only somewhat abstract may be at a more appropriate level of abstraction and hence more reliable.

Different Stages of Science

If a particular language of thought is correct in the sense that its structure is itself well validated, then it might be well be profitable to explore. This is the situation that prevails in what Kuhn (1962) called 'normal science' – a theory has been discovered and validated and then there is a stage of exploring the ramifications of this theory, applying the theory and using the theory as a means of guiding the search for new theories. This stage of science can be characterised as relatively cooperative and inward looking time – the participants tend to specialise into complementary skills and tasks and put these together within the established framework. There is a lot of 'building' on each other's work and the field establishes norms so that new entrants to the field are required to strongly situate their contribution within the established framework, for example by citing those considered authorities. This can have the effect of excluding outside ideas so that the field becomes inward looking. In extreme cases this results in the 'degenerate programmes' described in (Lakatos 1983).

During a period of normal science it may be sensible to simply accept the established principles, methods and assumptions and to concentrate on specialising and then developing complementary areas of knowledge using them. During such a time when those in the field are all using the same framework and outsiders are rare, one can take the common language of the participants for granted and simply use it as a vehicle for discussion.

Occasionally normal science is punctuated by periods of 'revolutionary science'. This is when the established framework (if any) has become (or is revealed as) unsatisfactory and if a new and better framework is introduces it may become accepted. During such a period very little can be taken for granted, especially the assumptions and methodology of the old framework. Instead of cooperation and complementarily, sharp competition between different ideas and methods dominates. Contributions are judged less by adherence to a particular framework and more by results. Typically in such periods one gets many contributions and academics from other fields being both offered and accepted.

During periods of revolutionary science one can not merely carry on with 'business as usual'. Contributions to knowledge need to be more thoroughly justified in terms of results and (since there is likely to be a diverse audience) explained without assuming that all will understand the same language of expression. Since even the framework is in flux, what the relevant authorities for citing are unclear and it is not necessarily helpful to use established methods.

Neither the simulation of MAS nor their design has an established and well validated framework. There is no ‘high theory’ of MAS, and no proven methods. Whilst it is true that some people have *claimed* the status of authorities, whether posterity will agree will depend upon how useful their contributions turn out to be. A paper that might well be acceptable by those inside a field during a period of normal science can be found wanting in periods of revolutionary science, especially in the extent to which it justifies its method and proves its usefulness through its results. (Edmonds 2000) discusses the relationship between formal systems and the dynamics of science in more detail.

A confusion about the stage that MAS is at may explain why some authors present their papers as they do – borrowing the style rather than the substantiality of papers in more successful sciences. If MAS did have a well validated general theoretical framework, then it might be more acceptable to present a exploration of part of that framework in a theoretical way, copying the methodology of accepted authorities in the field. Indeed, some of these papers *do* seem to imply that the use of simplistic deontic and epistemic logics *have* been established and proven, so what is left is to argue the details and make small extensions of these. Unfortunately this is far from the case – this style of formalism still has everything to prove.

What Sort of Logic is Suited for Modelling MAS?

Since, in common with many other styles of formal system, logic has the possibility of modelling any system (via the truths concerning that system), it not so much a question of whether logic per se is or is not the correct kind of system, but more the particular type of logic that is used[3]. In particular it has tended to be the axiomatics of non-temporal, context-independent and propositional logics which are commonly discussed in this domain. This is in keeping with the philosophical logic tradition briefly discussed above. However, it seems patently clear that, if one is going to use formal logics in this domain, that it is the formal semantics of temporal, and contextual predicate logics that are far more appropriate. I consider these aspects in the following subsections.

Time

There are many ways of interpreting what logic *is* – as many ways as there are of interpreting the syntactic systems that constitute formal logic. Some see it as a way of defining a set of truths using inference or formal semantics, others see the inference as the most important which can be used for inference of conclusions (including the set of truths). Different people put the emphasis on different parts (which they may see as primary) and see the other parts as coming from these. However you see it logic

[3] Although this still leaves question of the appropriateness of the implicit bias of the system.

relates a class of *truths* with a system of *inference*[4] (embodied either in the proof theory as allowable steps or as the formal semantic validity of expressions expressing an implication).

As such it is hard to see how a logic can usefully model the connection between goals and actions without including an explicit representation of time. For example, the relation between the goal indicated by the utterance "I want to go for a walk tomorrow" and the present action of "cancelling a meeting scheduled for tomorrow" has an important temporal element to it. Yet almost all of the logics that have to do with goals and actions (including the deontic and BDI logics) do not have any *explicit* temporal element, instead they attempt to capture either the instantaneous or unchanging aspects in the relationship between such as: beliefs, desires, norms, goals, actions. In the first case they must miss the dynamic nature of the relationships, for example that one might change one's intentions as the result of weighing the effect of violating a social norm – indeed such an approach rules out any *interaction* between these entities at all. In the later (unchanging) case, one is limited to modelling only those aspects of the relationship that are always the case – thus if sometimes (but not always) a belief changes a desire and sometimes (but not always) a desire changes a belief then these relationships will not be universal over time. In this case it is an implicit assumption that the important relationships *are* abstractable without reference to temporal contingencies, which is extremely unlikely and without justification[5].

The other approach is to use implication as an implicit model of causation and thus encode the relevant sequencing in the axioms. The result of trying to fudge the issue in this manner is that the essential elements of the situation are represented by ludicrous propositions such as *A = the assertion that state of the world is such that I will be walking tomorrow* and *B = the assertion that I will take an action which I believe will prevent a future event which would imply* $\neg A$. This sort of move does nothing to convince me that this method of formalisation is capturing the essence of the case. Yet this is the case with many attempts which attempt to concertina concepts which are temporality situated into a non-temporal framework – representing *processes* as *single states* is bound to lead to huge practical difficulties if the framework was ever used for real problems.

Lack of Formal Semantics

Another strange fact about the style of formal logic that have been discussed in RASTA and more generally in MAS is the lack of formal semantics. If one is primarily concerned with the *meaning* of modal operators and determining which ones are *valid* then the formal semantics are much more relevant than the axioms and proof theory. A logic that had as its universe of models (models in the logical sense) a set of MAS outcomes (i.e. the set of possible MAS states over time) and showed that certain expressions were logically validated w.r.t. these semantics, would be a useful development. On the other hand if one is more interested in inference (being

---

[4] For a thorough discussion of the nature of logic see (Gabbay 1994).

[5] I know of no attempts to justify such an assumption, rather the development of such logical formalisms seems to be on the basis that *any* caputuring of such mental entities is impressive and hence interesting, so it is felt that simple plausibility is sufficient to justify such explorations.

able to infer conclusions from premises) then the proof theory is more important (in this latter case, one would expect minimal discussion of the meaning of operators and a focussing on the useful and interesting inferences that can be obtained using the proof theory).

Context Dependency

The typical presentation of logic in MAS assumes and depends upon the fact that all the reasoning is done within a single context. Sometimes this is explicit, but more often it is left implicit and only indicated by the test problems (if any). This is very strange because reasoning about norms, goals, intentions, learning is only feasible if one can relate these to the contexts, for example intentions may involve action in several different contexts or involve explicitly effecting what the context is.

Numbers

A final area I will deal with is the ability of logic for understanding or designing MAS that does not allow for an adequate arithmetic. MAS in which numbers play no significant part are hard to find, but despite this most of the logics proposed rule out any sort of predicate logic in which such numbers could be defined. The reason for this is, presumably, because the introduction of arithmetic means that there can be no complete formalisation of truth, that is to say that there will be no method of proof that will be able to prove all the truths. This is due to Gödel's incompleteness theorem. However, the goal of completeness is simply inappropriate for almost all MAS – we are never going to be able to prove all an MAS's properties. Thus eliminating numbers to retain completeness is not sensible – it is a case of changing the problem to suit the tool.

Of course, a temporal contextual predicate logic with semantics that can capture multi-agent belief will not be such a clean simple system as those frequently discussed, but that is appropriate because most MAS are not clean simple systems! In this case simplicity is certainly not indicative of usefulness, let alone truth (Edmonds 2002). Some will argue is that they are deliberately abstracting away from the detail of time, context, and numbers in order to obtain a general theory, but the burden of proof is then surely upon them to show that they have done this successfully. Justifying such extreme abstraction on the basis of a few intuitions does not wash – the wish for the 'magic' shortcut is strong but can not be relied upon.

On the other hand, if proponents of such formalisms tried to use their constructions on real problems or to model real systems, the inadequacies and over-simplicity would be quickly revealed. If (as I suspect) there were no adequate work around that preserved the logic then this would be revealed and if there were it would be demonstrated how and in what way this formalism would work,

The Audience's Viewpoint

When presenting results there is an understandable wish on the part of the authors to concentrate on what they have *done*. However, for the audience it is more important to first of all judge whether the work is worth learning about or even applying. This is because they are bombarded with ideas people have had and systems they have designed – they are not short of ideas, but they do need help in deciding which ideas or systems to invest their time and effort in. Everybody feels convinced that their ideas or systems will work, otherwise they would not be presenting them. Similarly, everybody has some sort of thought train that lead them along the path they took, so everybody has *some* good reasons for doing what they did. Thus the presence of good reasons for doing something does not help an audience distinguish between different ideas or systems, more is needed.

One claim for formal logic made during the discussion at RASTA 2002 was that it aids communication because it allows one to be precise about ideas. That they allow one to be precise is true, formal systems (even if totally misguided) at least make for an unambiguous common referent. This is particularly attractive for disciplines which are bedevilled by different approaches, vagueness and misunderstandings with respect to their key terms. Precision is definitely a virtue, but it is not sufficient to ensure good communication. Whether formal logic does or does not aid communication is an empirical matter. Frankly, I doubt whether this was true for the audience we had at RASTA, for these logic papers are only accessible to the small minority who had sufficient familiarity with formal logic to be able to fluently 'read' it.

Even if there we assume that formal logic *did* aid communication between those who had suitable training, this still is insufficient to justify such a presentation. Being crystal clear in one's communication is no good if what is being communicated is not worth the effort. What was being communicated in some of these papers was simply unproved ideas and intuitions – directly comparable to specifications for systems that have not been implemented or otherwise tested.

Further the fact that the ideas and intuitions were expressed using formal logical expressions served to prevent the majority of the audience from evaluating them, leaving this evaluation to an 'in crowd' who are, on the whole, already sympathetic to the approach. It is almost certain that if I had not been there (being a person who is both critical and sufficiently knowledgeable of formal logic) there would have been no discussion about the worth of the formal logical approaches presented. Now I am sure that it was not the intention of the formalists to use their formalisms as a way of preventing criticism or ensuring acceptance, but this would have been the effect.

Thus a paper which does not provide any evidence for the usefulness of a formalism (apart from the reasons that lead the authors to invent or extend it) simply fails to satisfy the justified norms of scientific communication because it ignores the needs of the audience to evaluate the suggestions. Further, a formal system that has been used for solving a real problem or modelling a realistically scaled MAS will be greatly improved and be more likely to introduce genuinely new ideas. Intuitions are highly biased by the current *Zeitgeist* which is why rubbing them against a real problem is more likely to provide new input than simply more discussion between other academics immersed in the same *Zeitgeist*.

A Common Argument for Formalism

However, a logician (or mathematician or whatever) may object in the following manner: "the history of the development of formal systems has included many systems that would have failed on your criteria and yet turned out to be immensely useful later - are you not in danger of arguing against similar advances with such warnings?" My answer is fourfold.

? Earlier, we did not have the huge number of formal systems we have today, and in particular we did not have the general systems mentioned above. Today we are overwhelmed by choice in respect to formal systems – unless substantial advances are made in their organization all new systems will need to be substantially justified if their clutter is not to overwhelm us.

? There are proper domains for formal systems that are purely conceptual: philosophy or pure mathematics. Presenting a formal system elsewhere implies that it is relevant to the people in the domain in which it is being presented. If it really is relevant to them this needs to be demonstrated.

? Even in pure mathematics presentations or publications are required to justify themselves appropriate criteria - novelty, expressiveness and soundness are not enough (although the other criteria perform a weaker role than when they are applied elsewhere). For example, in the examination of a doctoral thesis in pure mathematics once the soundness of the work is deemed acceptable it is the importance, generality and relevance of the results that are discussed.

? The cost structure of the modelling enterprise has changed with the advent of cheap computational power. It used to be the case that it was expensive in both time and other resources to use and apply a formal theory, so that it was important to restrict which formalisms were available. Given that the extensive validation of the success of formal systems was impossible they had to be selected almost entirely on a priori grounds. Only in the fullness of time was it possible to judge their more general ease of use or utility of their conclusions. Now this situation has changed, so that the direct validational assessment of a formal system can be achieved with relative ease for relevant cases.

One can choose to judge a formal system by the criteria of pure mathematics (or logic) that is show the system has generality and inferential power by exhibiting theorems and proofs. One can choose to judge it as applied mathematics, whose criteria include problem solving ability and relevance by demonstrating its *use* in modelling systems. What is not acceptable is to fail to demonstrate that it succeeds by any kind of criteria. Some of the formalist papers in MAS fail in precisely this way, they excuse themselves of solving particular problems but also fail to exhibit and substantial theorems and proofs.

Some Suggestions for the Way Forward

It should be clear that I am not against the use of formal logics as a tool for understanding MAS *per se*, but against using them in unhelpful ways, namely as a language for philosophical discussion. Intuitions that are relatively unconstrained and unvalidated have a poor track-record when it comes to real applications and problems, and formalising these in relatively simple (and, I argued, inappropriate) logics does nothing to solve this basic problem. Simply following the form of a philosophical tradition is insufficient to justify the presentation of work – an audience rightly expects some conclusions in the form of results by which they can evaluate the ideas. Yet we do need somewhat abstract and precise models to improve our understanding, and logics are an expressive and flexible kind of formal system. So might be the way forward?

Before suggesting some steps we might take, I will describe our domain as I guess it is. I think that the study of social systems in general, and MAS in particular, will be more akin to biology than to physics and the production of MAS closer to stock breeding and ecological management than to traditional engineering[6]. I think that there will be hundreds of essentially different 'species' of MAS, all of which will need to be individually described studied rather than their being adequately covered by any easily accessible universal principles[7]. I think that there will not be any easy 'short cut' to useful high theory, and certainly not via vague intuitions expressed in formal logic. Thus I would the following based upon analogies with other sciences:

? The development of new ways of collecting data and observing MAS;
? A considerable period of descriptive modelling (i.e. less abstract modelling) so that we have a way to compare different MAS;
? A building up of complete chains made up of models at different levels of abstraction so that each are each clearly related (or relatable) to less abstract models;
? The insistence that any abstract or formal theory is treated with scepticism until it proves its worth – the more abstract it is the more is has to prove;
? That, nonetheless, we continue to try to build models whose level of abstraction is justified by (and judged by) the evidence;
? The rejection of papers that merely specify things based on single cases, intuitions and expressiveness because they are premature;
? That, nonetheless, the greatest variety of formal systems should be encouraged as possible members of a 'tool box' for MAS practioners and studiers (but only accepted after they have shown to be helpful in at least one real case).;
? That papers suggesting formal systems for helping design MAS should demonstrate that it is feasible to design an MAS that works using them;
? That papers suggesting formal systems for understanding MAS should show that they do, in fact, capture the phenomena they claim providing either a successful prediction or a credible explanation of that phenomena;
? That the field resists the temptation to retreat into formalism and philosophy when it substantial progress is difficult.

---

[6] Or, at least, to a traditional account of what traditional engineers do. Engineers, in practice, don't actually act as these accounts would suggest. A classic example of this is the neatly ordered elicit; analyse; design; implement; test cycle that software engineers are supposed to follow.

[7] After all a 2D cellural automata with extremely simple rules for each node can implement a full Turing machine and hence, in principle, *any* computation.

This is a more *pragmatic* and less ambitious approach than many academics have hoped for or will accept. They will continue to dream of inventing the 'magic bullet' that allows us to shortcut the large amount of messy empirical work that will be necessary and take us straight to powerful high theory (as, indeed, do I in moments of weakness). However I think this has more chance of producing useful knowledge and, *eventually*, useful theory. We will always continue to need some sort of abstractions to help us search, but until we have some well validated examples we need to stay as flexible as possible and stay suspicious of easy or prevalent intuitions.

A Concluding Exercise

Look through some of the papers in this (and similar) volumes. Does the 'conclusion' state what was done and why it was done but not state any results or conclusions (other than that the authors think it is the right way to do it)? Is there any way of *evaluating* what was done using the information in the paper? Is there any way of knowing when the techniques or ideas described in the paper would be useful to apply and when not? Have you been informed of anything except the present state of thought of the authors? If there were no results, did the system (either formal or software) help demonstrate or communicate the authors ideas effectively? Where those ideas so good to warrant presentation with no results?

One way of stripping bare the impressive effect that a formal logic imparts is to imagine the same sort of paper but using a simulation instead of a logic. If the paper was one where a simulation was described along with the reasons why it was so designed, but the simulation was not actually run and no results were shown, would it make a satisfactory paper? I think not.

References

Conte, R., Edmonds, B., Moss, S. and Sawyer, R. K. (2001). Sociology and Social Theory in Agent Based Social Simulation: A Symposium. *Computational and Mathematical Organization Theory*, **7**:183-205.

Edmonds, B. (2000) The Purpose and Place of Formal Systems in the Development of Science, CPM Report 00-75, MMU, UK. `http://cfpm.org/cpmrep75.html`

Edmonds, B. (2002) *Simplicity is Not Truth-Indicative*. CPM Report 02-00, MMU, 2002. `http://cfpm.org/cpmrep99.html`

Gabbay, D. M. (1993) *Classical vs. non-classical logics : the universality of classic logic*. Saarbrücken : Max-Planck-Instituit für Informatik.

Gabbay, D. M. (ed.) (1994) *What is a logical system?* Oxford: Clarendon Press.

Gärdenfors, P. (1988) *Knowledge in flux : modeling the dynamics of epistemic states*. Cambridge, MA; MIT Press.

Hopcroft, J. E. and Ullman, J. D. (2001) *Introduction to automata theory, languages, and computation*. Boston: Addison-Wesley.

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

Lakatos, I. (1983) *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.

Moss, S., Gaylard, H., Wallis, S. and Edmonds, B. (1998). SDML: A Multi-Agent Language for Organizational Modelling. *Computational and Mathematical Organization Theory*, **4**:43-69. `http://cfpm.org/cpmrep16.html`

Whitehead, A. N. and Russell, B. (1962) *Principia mathematica*. Cambridge: Cambridge University Press (originally published 1913).