![JASSS logo]

Jim Doran (2011)

Two Outline Models of Science: AMS and HAMS

Abstract

Two abstract and computational models of the long-term process of science are proposed: AMS and HAMS. An outline specification of each model is given and the relationship between them explained. AMS takes an Olympian ("artificial world") view of science and its processes. HAMS is simpler and relatively more abstract and comprises only a small set of core processes. A first implementation of HAMS is described. How AMS and HAMS might be validated and used in experimental investigations is considered including problems that might arise. Further work is proposed. A brief coda concerns a related model of science formulated from an idealist rather than a materialist perspective.

Keywords:
Computational Models of Science, Individual-Based Modelling, Scientific Method, Belief Systems, Belief Verification, Idealism

Introduction

1.1   Science may be defined informally as the set of processes by which a community of individuals uses reliable methods to obtain reliable understanding of itself and its environment. By "reliable understanding" is here meant insights that when followed up by action yield collective benefit in the evolutionary sense: individual and group survival in the face of natural hazards, increase in the community's numbers, and the acquisition of further reliable understanding. Reliable understanding in this sense may loosely be identified with *scientific knowledge*.

1.2   In this paper I approach the modelling and simulation of science processes from a long-term and abstract perspective, and suggest how we may model the acquisition by a community of reliable knowledge of its environment and of the ways in which that knowledge can be used. Thus this paper is not concerned with issues of government science policy, or university attitudes to science, or economic aspects of science. These are secondary. It is instead a matter of the perspective of prehistory: the hundred thousand years and more in which humanity (*Homo sapiens sapiens*) has built up a collective model of itself and of the world around it sufficient to obtain its present position of (seeming) dominance on planet Earth.

1.3   For several hundred years now it has been widely accepted that the accumulation of reliable knowledge requires going to the material world for observation, and devising situations in which beliefs about the behaviour of the world (including, of course, living entities therein) can be tested. See, for example, *De Magnete*, the influential work published by William Gilbert of Colchester in the reign of Queen Elizabeth I (Gilbert 1958).

1.4   Any macro-model of science should be capable of being used to address the phenomena that are prominent in discussion in the philosophy of science, for example, the nature of scientific knowledge, the observational scientific method, the notion of conjectures and refutation (Popper 1972), *reculer pour mieux sauter* episodes, and Kuhn's paradigms (Kuhn 1996).

1.5   Knowledge can loosely be defined as true belief. It follows that in the models now to be outlined, beliefs can heuristically be separated into those that have been verified and hence constitute knowledge and those that do not.

1.6   A particularly interesting question is this: where does the idea of science as a method of acquiring useful understanding originate? Is this idea itself of importance or is it just the practice that matters? Can we simulate on a computer the birth, life and death of the idea of experimental science?

Some Principles

2.1   The following principles are being applied to the development of a computational model of science:

- The desirability of an all-covering model in which scientific belief and scientific processes appear as an integrated component of society in its environment. The primary requirement on the model is not that it should capture the intricacies and subtleties or even content of science except in the broadest sense, but that it should have something persuasive and insightful to say about the science process as a whole viewed at an abstract level, most obviously the observed long-term exponential growth in scientific knowledge.
- The model must be agent-based. It must be grounded in representations of individuals, and furthermore, these individuals must possess, to at least some degree, relevant operational and cognitive abilities, notably the ability to sense the environment, to act upon the environment and to choose actions, and to maintain and update sets of beliefs and knowledge.
- Scientific belief must be distinguished from "ordinary" belief, for without such a distinction the model will not be about science at all. I shall assume that scientific belief is verified belief that is obtained by a process of systematic observation, measurement and reasoning.

AMS and its Major Components

3.1   AMS is an *Abstract Model of Science*. The aim is for AMS to be executable on a computer in the style of agent-based modelling, with major components as follows (compare Doran 2006):

- A computational *multi-agent system* that is embedded within and can interact in a constrained way with (e.g. movement and observation over a limited distance) a computer simulated environment. Individual agents have located presence within the environment and possess non-trivial cognition including decision-making.
- A *body of collective belief* and knowledge associated with the system of agents, which is built up by individual and collective observation, speculation, inference and action.

- *Social Interaction* between agents—forms of inter-agent communication, of agent cooperation and organisation and of role taking in collective action.
- An *environment*. Given the adopted perspective of human prehistory, it is natural that the environment should be spatial, 2D or 3D, and perhaps spherical, with simple properties such as different types of terrain.
- Asexual cloning of the agents, possibly with variation, providing evolutionary pressure.

Some AMS Design Details

4.1 In a little more detail the AMS design is as follows. Within the AMS environment there must be not only be the challenges to the agents that make it difficult for them to survive (e.g. acquisition of possibly mobile "energy" sources, avoidance of hazards, perhaps even defence against predators) but also useful insights that can be discovered. That is, there must be insights supporting survival for a form of experimental science to uncover—what elsewhere I have called a *potential technology* (Doran 1989).

4.2 Consider, for example, iron working as it entered the human technological repertoire more than three thousand years ago. The discovery that the combination of fire and a certain type of "rock", suitably manipulated, can provide iron weapons and tools must have seemed astounding to those who first made it. Similarly we can make available for discovery within the model combinations of environmental entities that yield new entities with new and useful properties (Doran 1989).

4.3 Agents correspond to individuals. There is a fixed finite set of basic actions that agents can execute. Within agents are relevant elements of cognition: observation, action, planning, memory, learning, conjecture, and belief representation (e.g. the CLARION cognitive architecture (Sun 2006) or the MIAP architecture (Doran 2010)). These may be programmed using standard algorithms and methods of artificial intelligence, for example, production systems. In general, different agents will have different "content" so that agents may be heterogeneous, with some more effective than others.

4.4 At the heart of the AMS model of science is the notion of a *belief*. Beliefs are regarded as recurring *patterns of sensa*. A belief is thus akin to a *sensory memory*. Computationally, beliefs are sets of tokens derived from the (simulated) sensory process, with associated relative timing. Note that there is available a fixed finite set of possible sensa.

4.5 A *belief* may be either a simple *association* or a *prediction*. The former links together sensa that repeatedly are simultaneously observed. The latter, in effect, specifies that one or more sensa will occur *after* certain other sensa have previously occurred.

4.6 Beliefs are thus derived from observations. However, they are also subject to a process of *abstraction* which discards sensa from a belief to form one or more new "higher level" beliefs. Thus beliefs can be at many levels of *abstraction* and can be *conjectural*. Beliefs can (also) be generated at random but are then, of course, very unlikely to be *true*.

4.7 Truth in a belief means that it reflects closely to the content and behaviour of the environment as observed. However, beliefs can fail to match this "reality" and be *misbeliefs* (Doran 1998).

4.8 Associated with any particular belief within an agent are certain *attributes*, notably the *strength* of the belief, and whether or not it is *verified* (see later).

4.9 Observations and beliefs must be capable of being transferred from agent to agent by directed communication, or by broadcast.

4.10 Agents must be able to cooperate for collective planning and collective action to mutual benefit. Collectively held beliefs may be defined as those held by a majority or many agents. Collective action requires collectively held beliefs and is mediated by relevant multiple-agent planning algorithms.

4.11 Agents may also be arranged into teams, with varying degrees of leadership and central control. Thus there may be organisational roles available for agents to adopt, and processes by which they come to adopt them.

AMS: Scientific Beliefs and Scientific Experiments

5.1 Fundamental activities of science are: observation, conjecture, prediction, experimentation and testing of conjectures, sharing of information, replication of experiments, and integration of world views. These activities relate both to individuals and to the collective of agents as a whole. Simplified forms of these activities can be modelled in terms of operations upon AMS beliefs as defined.

5.2 Thus some of an agent's beliefs may be called *scientific beliefs*. The difference between an ordinary belief and a scientific belief is that the latter is derived from a combination of particular forms of observation ("scientific data") or inter-agent communication, or inference.

5.3 *Scientific verification* means that the belief is strong, and has been repeatedly confirmed by observation rather than being merely an abstractive conjecture or random generated. Thus the essence of a scientific belief is that its accuracy has been repeatedly checked by observation in a reliable way. Naturally, observational conflicts reduce the strength of a belief.

5.4 A communicated belief that is labelled by a sending agent as verified will not necessarily be accepted as verified by the receiver. It is an accepted part of scientific method that mere authority is not always sufficient.

5.5 *Scientific experiments* are types of individual or collective action which create or test beliefs. More specifically, an experiment involves one or more agents in bringing about a situation in which the beliefs of interest (recall that beliefs are compounded of sensa) might be expected to occur. Planning by an individual agent or multiple agents is needed to set up of experiments, that is, to bring about test situations in which observations of interest may be made. This is compatible with the standard artificial intelligence concept of planning, which sees planning as a matter of combining action representations (here predictive beliefs) to form a plan that when executed will bring about desired situations. Of course, there will need to be other *ad hoc* goal and planning management algorithms, part of the fixed processing of an agent, that determine when experiments should be performed and with what specific objectives.

AMS Implementation

6.1 In spite of the obvious complexity of AMS, there is little doubt that a version of it can be implemented. AMS does not have to be as complex as its natural language description, which is necessarily ambiguous, might at first suggest. *Simple* forms of everything mentioned in the foregoing specification are programmable. It is enough to capture in abstract the essential structures, distinctions and processes.

6.2 Little need be said here about the required computational infrastructure. A variety of agent software platforms are available. Indeed, it is now possible to run systems of tens of thousands of agents on high performance computers, for example by use of the FLAME framework developed jointly at the UK Rutherford Appleton Laboratory and the University of Sheffield (Greenough 2010).

AMS Validation

7.1 Standard modelling procedure requires validation of the model. Put at its simplest, there are two aspects to validation: the specification of the model's detailed structure to be consistent with reality to the maximum degree possible, and the testing of the completed model by comparing with observation the results it delivers.

7.2 One possibility is to validate AMS against particular strands of science and technology, for example, against the long-term development of iron working referred to earlier.

However, more relevant and compelling would be to demonstrate in the model the macro-behaviour of science that occurs in reality. Thus *AMS (and HAMS, see later) must be such that realistic macro-behaviour is emergent.*

Experimenting with AMS

8.1   The AMS model can be run starting with the agents initially having no scientific beliefs and, assuming that the environment provides opportunities for discovery, it will be possible to observe a body of collective scientific beliefs emerge within the agent community over one or many generations. Recall that within the AMS framework, sets of beliefs, whether or not they are scientific, will be subject to the evolutionary pressure of competition for resources between the agents.

8.2   We might hope to discover insights about the emergence of a body of scientific knowledge ("true" beliefs), and its dependence upon a range of factors, for example, the degree of agent heterogeneity, the level of agent interaction, and the types of organisation that the agents can support. In particular, we might hope to demonstrate how having knowledge leads to the further acquisition of knowledge. This can happen because, for example, having knowledge "frees up time" from survival activities for "scientific exploration" by agents and also makes that exploration more productive.

8.3   AMS as described in outline is certainly not a single model. There are many possible structural variations left open as well as potential adjustable parameters. Therefore a systematic study of the structure variation and parameter space is essential, looking for performance and emergent phenomena of interest. But the parameter and structure variation space is both probabilistic and very large.

8.4   In principle such models as AMS are finite and can support precise notions of optimality. Thus it may be possible to define and determine *analytically* (i) the maximum (in a defined sense) body of scientific knowledge (i.e. set of "true" beliefs) that a community of agents in the model can possible share for a particular environment, and also (ii) to determine whether that maximum set of beliefs can actually be reached by the agents.

From AMS to HAMS

9.1   AMS seems in danger of accumulating a vast and intractable weight of detail. Yet AMS components and their interaction must be formulated to be to some degree realistic if the model is to be of more than of purely computational interest. Whilst it is easy to specify such processes as perception, abstraction, collective planning and organisational role-taking as being part of a model, the history of artificial intelligence studies demonstrates that such processes are always difficult to program non-trivially on a computer, especially if they are to interact coherently.

9.2   To model science a nicely judged set of design decisions must be made that steer a course between the Scylla of overwhelming detail and the Charybdis of over-abstract irrelevance. Arguably AMS retains considerable structure and detail that is inessential for the phenomena we wish to address. So now I offer a simpler, more abstract model of science: HAMS (a *Highly Abstract Model of Science*) that steers a course significantly closer to Charybdis.

HAMS in Outline

10.1   HAMS is built around *agents*, each associated with a set of *propositions*. Propositions may merely be *held* by an agent (i.e. without being believed), or may be *believed,* or may be *verified*. Intuitively they abstractly correspond to a plethora of ideas such as: "Ostriches bury their heads in the sand", "White light can be split into many colours", "The Earth we live on is spherical" and "Humans and chimpanzees have common ancestors". However, within HAMS *propositions have no internal structure.* Nevertheless, some sets of propositions *conflict*, meaning that all the propositions in the set cannot simultaneously be believed by an agent.

10.2   There are means by which agents generate propositions to be *held*, and means by which propositions held by an agent may come to be *believed*. Subject to *environmental verification test preconditions* (e.g. that certain other propositions are believed) a proposition may be *tested* by an agent to determine if it is to become (objectively) *verified*. Verification is by reference to a global *environmental look-up table* by which, in effect, environmental properties are captured.

10.3   Agents *pass* propositions amongst themselves, along with their belief and verification statuses, in accordance with a *communication network*.

10.4   Subjecting held or believed propositions to verification tests is to be interpreted as *science*. Scientific models and theories are thus combinations of propositions. Non-science, by contrast, is a matter of holding believed propositions without verification.

Detailed Formulation of HAMS

11.1   There is a set of *agents*, and with every agent an associated (dynamic) set of *propositions*.

Every possible proposition (a large but finite set) has an associated global *environmental truth value* accessible by lookup in an *environmental {P→T/F} table*. There is no environment as such.

A proposition associated with an agent has one of the statuses: *unknown, held, believed* or *believed and verified*. "Unknown" implies that the agent has no awareness of the proposition at all.

Agents are the nodes of a *communication network* and pass propositions randomly.

A proposition that is passed by one agent to another passes its status unchanged. Where this leads to a clash within the receiving agent, this is resolved, perhaps randomly, within that agent.

There are *association rules* that generate new propositions for an agent to hold from the propositions it currently holds.

Entailment rules apply to held propositions and generated new believed propositions. Intuitively, entailment rules express both "reasoning" and "observation".

Rules have the structure {P1 & P2& … ⇒ P}.

All propositions generated by simple observation (i.e. by an entailment rule with empty LHS) are both believed and verified and have environmental value T. Typically not all propositions can be obtained by observation.

All propositions generated by an entailment rule from a set of believed propositions are believed.

An *environmental verification testing precondition* is a set of believed propositions all of which must be believed (but not necessarily verified). A believed proposition (that has not already been tested) may be tested for verification when its corresponding precondition is satisfied.

When a proposition is tested it becomes verified iff its environmental value is T.

A proposition that is held but not believed (and so is *a fortiori* not verified) may come to be believed in the following ways:

- Randomly
- By receipt of it as believed from another agent
- By execution of an entailment rule

A proposition that is held and believed but not verified may come to be held but not believed in the following ways:

- Randomly
- By receipt of it as not believed from another agent
- By being tested but NOT verified

A proposition that is held, believed and verified within a particular agent may change its status if it is received with a different status from another agent.

There is a set of *conflict sets of propositions* that cannot all be believed simultaneously by an agent. Whenever an agent comes to believe complete conflict set of propositions, its belief in at least one of them must be (perhaps randomly) cancelled.

The proposition conflict sets must be consistent with the environmental {P→ T/F} table, that is, conflicting propositions cannot all have environmental value T.

HAMS Processing

```
Initialise environment P → T/F table
Initialise proposition conflict sets
Initialise agents and with their associated proposition sets (all with status held)
Repeat {
 For each agent: execute (randomly) selected matching association rules
 For each agent: execute (randomly) selected matching entailment rules
 For each agent: process communication, including handling clashes with incoming information
 For each agent: where preconditions are satisfied, (randomly) perform environmental verification tests and
  update agent proposition sets and statuses correspondingly
 For each agent: (randomly) resolve any proposition belief conflicts
}
```

11.2   The foregoing specification constitutes the HAMS model. Clearly HAMS is highly abstract. "Reasoning", for example, is expressed as an arbitrary set of entailment rules. The different types of rules and the test preconditions may be generated randomly or in accordance with some structural framework selected by the experimenter for a particular experiment. Similarly, rules may be selected for execution randomly and conflicts between proposition statuses within agents may be resolved either randomly or biased in particular ways.

Implementation of HAMS

12.1   The specification of the previous section is sufficiently precise for it to be relatively straightforward to implement a version of HAMS in any general purpose computer language. The main implementation tasks are choice of data structures—easy—and, much more importantly, the precise form of key pseudo-random based events, for example, what exactly happens when proposition statuses conflict. The latter imply a range of adjustable parameters and structural alternatives upon which the behaviour of the system depends.

12.2   There exists (June 2011) an initial implementation of HAMS in the programming language C, which has verified computational completeness and coherence.

Experimenting with HAMS

13.1   What can we reasonably hope to learn from HAMS? Returning to the major phenomena of science mentioned at the outset, we might hope to conduct experiments that demonstrate exponential growth in the set of verified propositions (to be interpreted as growth of scientific knowledge), to demonstrate processes by which distinct parallel systems of propositions emerge one (or more?) of which is formed of verified propositions, and to demonstrate the HAMS equivalent of episodes of *reculer pour mieux sauter*. More remote, perhaps, is the possibility of demonstrating a HAMS analogue of a shift from one scientific paradigm (Kuhn 1996) to another.

13.2   Sensitivity analysis by way of systematic experimental trials is essential to determine under what circumstances, that is where in the parameter and micro-structure alternative space, particular types of macro-behaviour emerge.

13.3   It should be noted that HAMS is sufficiently abstract and simple in its structure (at least compared with AMS) that insights into its behaviour by direct analysis rather than by computer experimentation seem possible. This is under investigation. Furthermore, quite aside from its role as a model, HAMS seems to be an interesting computational process in its own right.

General Discussion

14.1   Both AMS and HAMS have the potential to yield significant advances in our understanding of science processes, partly *because* of the high level of abstraction at which they are pitched. However, thoroughly to implement and test either of them is a major task.

14.2   These two models are at different levels of abstraction. Which of them would be the more informative in practice? As yet agent-based modelling method cannot answer such questions prior to actual experimentation. However it is clear that the two different models address different types of question. For example, a core issue in any such model is the possible content of the propositions it manipulates and how the propositions are handled. AMS, as it is so far specified, assumes that propositions are structured as certain combinations of sensa. HAMS says nothing directly as to the possible content of propositions but recognises the existence of association and entailment in a very unstructured way, and also recognises conflicts between propositions. This means that AMS can, in principle, address issues concerning the structural similarity between propositions and the way in which propositional structure can determine association and entailment. HAMS cannot. Incidentally, it is tempting to adopt a more formal approach to knowledge representation and to try, for example, to deploy first order mathematical logic. But history strongly suggests this would be to enter a laborious dead end.

14.3   A related point is that the four possible statuses currently recognised in HAMS (unknown, held, believed, and verified) could certainly be developed into a more elaborate range, but not necessarily with overall benefit.

14.4   Can the *idea* of experimental science be expressed as a proposition in AMS and/or HAMS and hence emerge and be used by the same model algorithms that handle all propositions? If so, this opens the door to the model being used to address the *discovery* of scientific method, not merely its use. In AMS, building a proposition that is descriptive of the scientific method looks very difficult. In HAMS, however, it would be possible to designate one proposition, call it P#, and so select association, entailment and environmental verification test precondition rules that (i) P# does not easily come to be held or believed, and (ii) that belief in P# is a precondition for verification of the majority of propositions. This would make P# a *sine qua non* for most verification, and therefore for the large-scale development of scientific knowledge as it is modelled within HAMS. In an experimental trial of HAMS one would expect to see a time at which P# comes to be believed quickly followed by a surge in verified beliefs. The verification of P# itself is more obscure. Interestingly, P# need not be selected by the experimenter. If the rule sets are generated at random, then a proposition (or more than one proposition) may play the role of P# merely by chance.

14.5   Finally, what can be done with the attractive idea that the community of scientists is itself a social unit and therefore appropriately modelled as a single agent? It seems relevant that in HAMS the various sets of rules (association and entailment) can be common to all agents or, clearly, they can vary from agent to agent. More generally,

it may sometimes be possible to replace a set of agents by a single agent, without change in overall system behaviour, if the intra- and inter-agent processes are suitably adjusted. The downside is likely to be unnatural and potentially intractable complexity.

🌐 Conclusions

15.1 Both AMS and HAMS seem sufficiently promising in the interest and relevance of their behaviour to merit further attention. This is partly because of the inherent interest of the complex computational processes they embody, but it is also, of course, because of their relevance to actual science, its history, present state, and future development. The importance of identifying previously unrecognised core processes and relationships within science, however abstractly formulated, is a goal worth striving for.

🌐 Coda: I-HAMS, an Idealist Alternative

16.1 The models developed in this paper are open to challenge in a fundamental way. They assume that in reality there is a material environment to be modelled, or more subtly, that it is correct to model *as if* there is a material environment. This seems natural enough, at least within the context of a (perhaps naïve) computer science and/or artificial intelligence metaphysics. Materialism is, after all, at the heart of the AI endeavour. But what if this is *not* the case in reality? What if, say, some form of dualism is out there? More radically, what if reality, including that part of it which is to be scientifically described, does not exist as *matter* waiting for us to discover it, but is instead a *collective mental construct* of those that experience it? Consider the cautious remark: "We are accustomed to regard as real those sense perceptions which are common to different individuals, and which therefore are, in a measure, impersonal" (Einstein 1956, page 2).

16.2 It might seem that such idealism (analysed, for example, by Bishop Berkeley 1710) precludes computer modelling. But this is not so. Indeed, there is nothing to stop us creating a model, I-HAMS say, in which the agents formulate conjectures, potentially including scientific conjectures, that duly turn out to be sound *iff a sufficient number of agents adopt them*. For a description of an implemented version of such a model see Doran (in preparation). This can lead to stable collective belief systems, and may perhaps be seen as a strong form of what is sometimes called the *social construction of reality*.

🌐 References

BERKELEY, G. (1710). *A Treatise Concerning the Principles of Human Knowledge*. Dublin.

DORAN, J. E. (1989). Distributed AI based modelling of the emergence of social complexity. *Science and Archaeology*, 31: 3-11 (Special volume comprising computer archaeology papers presented at the Dynamic Text Conference, Toronto, June 1989).

DORAN, J. E. (1998). Simulating Collective Misbelief. *Journal of Artificial Societies and Social Simulation*, 1 (1) 3 http://jasss.soc.surrey.ac.uk/1/1/3.html.

DORAN, J. E. (2006). Agent Design for Agent Based Modelling. In F. C. Billari, T. Fent, A. Prskawetz, J. Scheffran (Eds.), *Agent Based Computational Modelling: Applications in Demography, Social, Economic and Environmental Sciences* (pp. 215-223). Heidelberg: Physica-Verlag (Springer). [doi:10.1007/3-7908-1721-x_11]

DORAN, J. E. (2010). The MIAP Cognitive Agent Architecture and its Potential Use in Agent-Based Economic Models. Paper read to *Advances in Agent-Based Computational Economics* (*ADACE 2010) Conference,* July 2010, Bielefeld, Germany. (Text available from author).

DORAN, J. E. (In preparation). Computational Models of the Nexus between Time, Agents and the Material World, (Unpublished complete draft available from the author).

EINSTEIN, A. (1956). *The Meaning of Relativity* (6th ed.) London: Methuen. (1st edition 1922).

GILBERT, William (1958). *De Magnete*, (Tr. P. Fleury Mottelay). New York: Dover Publications Inc. (First edition in Latin published by Peter Short, London, 1600)

GREENOUGH, C. (2010). Parallel Implementation of Large Scale Agent-Based Models in Economics. Paper read to *Advances in Agent-Based Computational Economics* (*ADACE 2010) Conference,* July 2010, Bielefeld, Germany.

KUHN, T. S. (1996). *The Structure of Scientific Revolutions*. Chicago and London: University of Chicago Press (3rd edition). [doi:10.7208/chicago/9780226458106.001.0001]

POPPER, K. (1972). *Conjectures and Refutations: the Growth of Scientific Knowledge*. London: Routledge and Keegan Paul (4th edition, revised).

SUN, R. (2006). The CLARION Cognitive Architecture: Extending Cognitive Modeling to Social Simulation. In R. Sun (Ed.), *Cognition and Multi-Agent Interaction: from Cognitive Modelling to Social Simulation* (pp. 79 - 99). Cambridge: Cambridge University Press.