Manchester Metropolitan University Business School Working Paper Series

Bruce Edmonds
Centre for Policy Modelling

Artificial Science – A Simulation to Study the Social Processes of Science

05-06

September 2005

ISSN 1478-8209

The Manchester Metropolitan University Business School is one of the largest business schools in the UK comprising more than 150 academic staff organised into nine thematic subject groups. The Working Paper Series brings together research in progress from across the Business School for publication to a wider audience and to facilitate discussion. Working Papers are subject to peer review process.
The Graduate School of Business of the Manchester Metropolitan University publishes management and business working papers. The Graduate School of Business is the centre for post-graduate research in all of the major areas of management and business. For further information contact: The Director, Graduate School of Business, Manchester Metropolitan University, Aytoun Building, Aytoun Street, Manchester M1 3GH
Telephone No: 0161 247-6798. Fax No 0161 247 6854

Introduction

Science is a collective enterprise – it is not simply the aggregated efforts of individuals. In fact, some writers (e.g. [1]) go so far as to claim that the social processes special to science are the *only* thing that distinguishes it from other activities. In any case the social processes are critical to the success and character of what we know of as science. Here I exhibit a simulation that explores some of these.

Traditionally there is the 'building-block' picture of science [2] where knowledge is slowly built up, brick by brick, as a result of reliable contributions to knowledge – each contribution standing upon its predecessors. Here, as long as each contribution is checked as completely reliable, the process can continue until an indefinitely high edifice of interdependent knowledge has been constructed. However other pictures have been proposed. Kuhn in [3] suggested that often science progresses not gradually but in revolutions, where past structures are torn down and completely new ones built.

Despite the importance of the social processes in science to society, they are relatively little studied. The philosophy of science has debated, at some length, the epistemological aspects of science – how knowledge is created and checked 'at the coal face of the individual'. Social processes have been introduced mainly by critics of science – to point out that because science progresses through social processes it is 'only' a social construction, and thus has no special status or unique reliability.

Here I take a neutral view, that is it is likely that there are many different social processes occurring in different parts of science and at different times, and that these processes will impact upon the nature, quality and quantity of the knowledge that is produced. It seems clear to me that sometimes the social processes act to increase the reliability of knowledge (such as when there is a tradition of independently reproducing experiments) but sometimes does the opposite (when a closed clique act to perpetuate itself by reducing opportunity for criticism). Simulation can perform a valuable role here by providing and refining possible linkages between the kinds of social processes and its results in terms of knowledge. Earlier simulations of this sort include Gilbert in [4]. The simulation described herein aims to progress this work with a more structural and descriptive approach, that relates what is done by individuals and journals and what collectively results in terms of the overall process.

The Simulation

The General Structure

The simulation involves a fixed number of agents (representing individual or closely collaborating teams of scientists) a journal (only one in the present simulation) which includes the set of formal sentences representing the knowledge that is discovered and published. Each agent has a private store of knowledge which may or may not be public (i.e. an axiom or published) – this store is their working knowledge. To use a public item of knowledge this must be added to their private store before they can use it to produce new items. They submit some of this to the journal which selects (according to some criteria) a subset which is then published and becomes available to others. The whole set-up is illustrated in Figure 1.

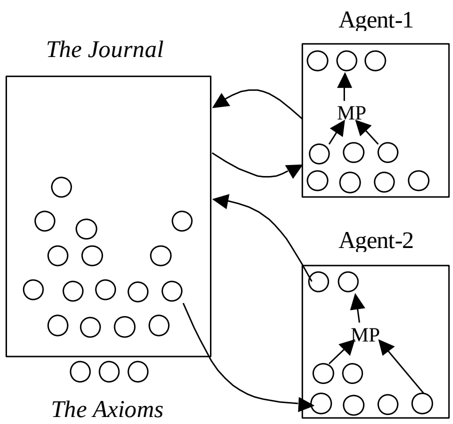

Figure 1. An illustration of the set-up with two agents (circles are items of knowledge, rectangles are agents)

The Environment and Task

Science continually progresses into the unknown. In science sometimes the end points are known – for example, when it is known that a certain disease is passed on genetically, then the genes that are responsible may be sought. Often, however, scientific discoveries are a surprise to their discoverers. Thus it is often the case that scientists do not know exactly what it is they are searching for. This is in contrast to engineering where it is usual to know the problem for which an answer is sought.

This poses a problem for a would-be simulator of the social and cognitive processes that contribute to science – *how can one simulate creative discovery of the unknown*?

The answer I have chosen is to use a formal system (logic) as the representation of knowledge, so that the agents work on the logical structures to produce new structures (theorems in the logical sense), but where it is impossible to know in advance how useful these will be. This decision has distinct consequences both in terms of the possibilities and limitations of the model and in terms of the assumptions on which it relies. These will be discussed later. This can be seen as following [5].

Thus the universe of knowledge that the agents will explore in this simulation is the set of inferable formal sentences derivable from a given set of initial axioms. For ease of implementation I have restricted my self to logics formalisable as Hilbert Systems (that is, ones with a set of axioms and a single rule of inference, Modus Ponens, which is recursively applied, see an introduction to logic, e.g. [6]). The agents can produce new sentences by applying existing sentences to other sentences using Modus Ponens (MP). The form of this is if you know A and you know $A \to B$ then you can also conclude B (written $A, A \to B \models B$). An example of this is: when A is $((a \to a) \to (a \to a)$ and B is $(a \to a)$: from $((a \to a) \to (a \to a)$ and $((a \to a) \to b) \to b$ we can infer $(a \to a)$. This is illustrated in Figure 2.

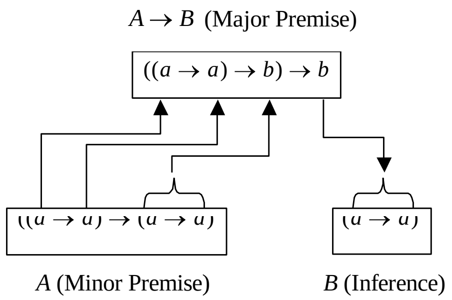

Figure 2. An illustration of the working of MP

The agents thus have the task of discovering new formal sentences. The advantages of this structure are that: (1) future developments from any item of knowledge are not known in advance; (2) knowledge is not only useful as an end in itself but can be used as a tool to act upon other knowledge to produce new knowledge (as the major premise in MP, the *A* in Figure 2); (3) the programmer of the simulation does not necessarily know *how* one gets to any given theorem of the system, which reduces the temptation to bias the simulation to get specific results; (4) the task is suitably *hard*, as the development of automatic theorem-provers shows.

In order to set up the field of knowledge that the agents will collectively explore the simulator needs to list the symbols being used and list the axioms of the relevant logic. Optionally the simulator can also list a number of known theorems that are considered important by logicians and give them a value, though *how* one derives these is not needed to be specified (this is for the agents to find out). These 'target theorems' are unknown to the agents until they discover them. They represent (in the loosest possible way) useful technologies that may come out of science. Counting how many of these have been discovered (and the total value of their 'worth') is an indication of the effectiveness of the collective discover effort, and can be a better measure that simply counting how many new sentences have been discovered since it is easy to develop trivial elaborations of already known sentences.

The Agents

In this simulation the agents have a very simple-minded approach to the production of new knowledge: agents select two items in its own store of knowledge and apply the MP rule to it, which may or may not result in a new item of knowledge which is added to their store. Each agent has two private stores of knowledge: *firstly*, a store of formal sentences that are candidates as the minor premises for the MP rule and *secondly*, store composed of candidates for the major premises. The former roughly corresponds to the primary knowledge of a scientist and the second as the set of techniques of the agent since it determines which transformations can be applied to which items and what can be produced.

Each time period the agent does the following:

1. Decide what new items of knowledge (both major and minor) to add to its private store from the published set, also which to drop.

2. Decide which major premise and what set of minor premises it will try with the MP rule and add any results to its (minor) store.

3. Decide which of its private knowledge (that is not already public) it will submit to the journal.

There are obviously many different ways of making these decisions. Each of these ways will have a (varying) impact upon the development of the collective knowledge. In addition to the above, gradual, update policy if the agent fails to discover any new sentences during a given number of consecutive time periods it may 'panic' and completely replace one of its stores with a new set of sentences.

Key parameters and setting of the agent are as follows. For each of its private knowledge stores (minor and major) the update policy includes the following: its size; the rate at which it adds or drops knowledge from this store; *how* it does either the addition; the dropping; or the panic replacement (at random/probabilistically the best/the best judge either on raw past fertility or past fertility with a bias towards simplicity); whether it panics and how long it endures lack of success before panicking; which to try (the best/probabilistically the best/untried/random); and how it judges what it knows (personal fertility/lack of failure to produce new knowledge).

Also its submission policy: whether it submits all novel (i.e. unpublished) sentences to the journal or only the simplest/best ones.

The Journal

The journal (*the Journal of Artificial Sentences and Successful Syllogisms*) is the gatekeeper to the repository of public knowledge. The key aspect of the journal is the criteria it uses for assessing the items submitted to it, so as to decided what (if any) it will publish. This occurs in three basic stages: the short-listing of those that met basic criteria; the evaluation of those short listed; and their ranking. The journal then published a selection of the top *n* in the ranking (if there were more than *n* short listed), otherwise all of them. This final selection could be the best (from the top); probabilistically on the weighted score (the higher the score the more likely it is to be selected); randomly or simply all of them. The evaluation of the submissions was done as a weighted sum of scores for a number of aspects: the number of variables in the sentence, its brevity, the extent to which it shortens sentences when used in MP, and the past success of the author. The weights and selection policies can be set by the programmer.

Methods of evaluation

Key to many of the decisions made by the agents or the journal is the evaluation of the existing knowledge. Ultimately this can be considered as a guess at the future usefulness of that knowledge, in terms of either: its productivity in producing new knowledge; reaching the hidden 'target theorems'; or in getting published. This may be done in a number of ways. One way is by looking at the historical record of how productive the sentence has been in the past in resulting in new published knowledge (this can be done in a recursive way to value sentences that have produced productive sentences etc.). Another way is to look at the most published agents and see what knowledge they have used (in published work). Other ways include considering features of the sentences themselves, for example measures of their simplicity (how many variables they have, how long they are, to what extent the sentence results in a shortening of sentences when applied using MP, etc.)

Preliminary Results

At the time of writing only preliminary results are available, which explore only a very small proportion of the possibilities inherent in this model. By the time the full paper is due I expect to have a better feel for the nature of some of the results and the deeper limitations of the model structure. However a summary of the indications so far obtained follows.

Many of the settings do affect the outcomes to a significant degree. However many which increase the short-term success (measured in a number of different ways) of the scientific progress also have the effect of reducing the longer-term maintenance of new results. Thus, for example, adding new sentences *at random* to an agent's private knowledge (i.e. regardless of the agent's evaluation of sentences) decreased the short-term level of discovery markedly, but then that level of discovery lasted a longer time. In contrast where agents follow other agents closely (preferentially adding sentences used successfully by others) results followed much more quickly to begin with but then petered out to zero after 40-60 time periods (only then deviating from zero when an agent panicked and hit lucky with its new set of knowledge). Such a result would indicate that a process of fairly frequent, but collective revolution was one of the most efficient collective modes of discovery.

In general most of the targeted sentences were either discovered very soon, or never. This suggests that "deep" sentences (those difficult to reach in this collective but individually stupid manner) require guidance from a deeper knowledge of the individual logics concerned, and is not so amenable to a generic approach (collective or otherwise).

Discussion

The Possibility of Limited Validation

Following [4] it may be possible to compare the structure of the published knowledge that results in this simulation (i.e. which authors/items are derived from which previous items by which authors) might be compared with the structure found in citation indexes such ISI using a number of measures, statistics or structural comparisons. Unfortunately negotiations with ISI indicate that they are only prepared to part with the structural information of their databases (suitably anonymised) for rather large quantities of money (i.e. around $30000). If anyone knows of an alternative source, please contact the author.

Limitation and Extensions

Clearly many of the limitations in this simulation are arbitrary: Thus I list a few possible extensions as examples:

- decision methods of arbitrary complexity can be implemented in agents (indeed these methods could themselves be evolved by GP);
- there could be many journals so that the prestige of a journal; its impacts and the quality of its submissions could be allowed to develop with the simulation;
- instead of inferring new knowledge the agents could hypothesise and test candidate sentences performing tests on the logical semantics (e.g. a row of the truth tables in classical logic);
- a peer review system could be implemented whereby reviewers are selected depending on their past publishing success and impact; they could use their own experience of what is useful as their criteria for judging entries and their own tests; and items could be selected resulting on the votes of reviewers;
- informal social networks could be introduce to pass knowledge from agent to agent other than via official journals;

- agents should be allowed to reproduce in terms of the students they teach and retire after a suitable time (or if they are spectacularly unsuccessful).

More fundamentally the present structure of the simulation assumes that there *is* some independent ‘correct’ knowledge to be discovered and that it is checkable. This could be corrected by providing some database of atomic facts (e.g. the linkage structure of part of the web) and then hypotheses about these could be attempted to be induced (as in inductive data-mining techniques). The journal (or journals) would not be able to 100% check the veracity of any knowledge but have to rely on some fallible process to come to a judgement upon the knowledge. However, a disadvantage of such an approach is that it would lack the tight inter-dependency of knowledge that seems to be such a characteristic of some sciences[1].

Relationship with Distributed Theorem Proving (DTP)

The simulation *is* a forward-chaining theorem prover, and can be seen as an answer to [7] since it could be truly distributed. However it is a very inefficient one – it is deliberately *generic* in that it has not been ‘tuned’ for performance (by using deep properties of the particular logic being investigated), since this is not its goal. Despite this, lessons learned in this simulation do have potential in terms of informing the design of distributed theorem provers and *vice versa* from what is discovered about efficient DTP to this simulation (and potentially science itself[2]).

OTTER [8], a particular and quite successful theorem prover is quite close to the how a single agent works in the above simulation. It has a list of candidate minor and major premises and works on these to extend the set of known sentences until it reaches the target theorems. It allows for a large range of techniques in re-writing formulas, guiding search and applying rules that are not touched upon here.

Conclusion

I hope to have shown how it is possible to capture some aspects of the social processes that contribute to the construction of science. Such modelling has the potential to intermediate between observations concerning how science works and

[1] Of course it *may* be that this *is* more appropriate for the social sciences.
[2] One can but dream!

areas of distributed knowledge discovery in computer science, e.g. automated theorem proving. It could help sort out the roles of the different processes in science confirming or disconfirming philosophical speculations (such as [9]).

References

[1] Longino, H. (1990) *Science as Social Knowledge*. Princeton: Princeton University Press.

[2] Hempel, C. G. (1966) *Philosophy of natural science*. Englewood Cliffs, N.J. :Prentice-Hall.

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

[4] Gilbert, N., Pyka, A. and Ahrweiler, P. (2001). Innovation Networks - A Simulation Approach. Journal of Artificial Societies and Social Simulation vol. 4(3) <http://www.soc.surrey.ac.uk/JASSS/4/3/8.html>

[5] Langley, P., Simon, H.A., Bradshaw, G.L. and Zytkow, J.M.. (1987) *Scientific Discovery: Computational Explorations of the Creative Processes*. MIT Press.

[6] [copi introduction to logic]

[7] Fisher, M. and Wooldridge, M. (1977) Distributed Problem-Solving as Concurrent Theorem-Proving. In Boman and van de Velde (eds.) *Multi-Agent Rationality*, Springer-Verlag.

[8] L. Wos, R. Overleek, E. Lusk, and J. Boyle (1984) *Automated Reasoning: Introduction and Applications*. Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[9] 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)