Modelling Intervention Strategies for Cooperative Environmental Management

Jim Doran

Modelling Intervention Strategies for Cooperative Environmental Management

Jim Doran

Department of Computer Science, University of Essex, Colchester, CO4 3SQ, UK
(doraj@essex.ac.uk)

Abstract: Central to sustainable natural resource management is the achievement of cooperation and collective action amongst stakeholders with initially conflicting short and long-term goals. It is argued that automatically generated agent-based computer models may be used to explore the ways in which external intervention can bring about effective stakeholder cooperation in environmental resource management contexts. The potential advantages of the agent-based modelling approach in this context include objectivity, and the discovery of currently unrecognised intervention strategies of practical value. An experimental procedure is proposed, and, by reference to a detailed design for a class of agent-based models, the technical obstacles that must be surmounted before this potential can be realised are examined. They include combinatorial complexity, and difficulty in the interpretation of a model's behaviour in human social terms.

Keywords: multi-agent systems, agent-based modelling, environmental modelling, intervention strategies, watershed management

1. INTRODUCTION

Based on multi-agent systems (MAS) theory and practice [Weiss, 1999], computer and agent-based modelling of social and organisational systems [Doran, 1997, 2001] is becoming of practical value in a range of application domains [Moss and Davidsson, 2001] including the environmental.

Here we view a *multi-agent system* as an interacting and probably inter-communicating collection of agents sharing a common (possibly simulated) environment, where an *agent* may loosely be viewed as an "object" in the software engineering sense that possesses a degree of autonomy and cognitive ability. This last requirement indicates the relevance of artificial intelligence [Russell and Norvig, 1995].

Cooperation is a key topic in MAS studies [Doran et al., 1997]. We define it here as intentionally coordinated action. Most agent work on cooperation concerns how to design cooperation into a MAS, or how to model existing cooperation, rather than how to achieve it in a pre-existing non-cooperating set of agents. But *achieving cooperation* in a pre-existing situation is very often the key real-world problem.

The importance of agent-based modelling on a computer lies in its ability to make discoveries otherwise impossible. Empirical observation and abstract theorising has its limits. New and important insight can be obtained by the use of computational rigour in the form of computational models. Models based upon agent cooperation are required in order to model the impact of complex social actors.

2. AN ENVIRONMENTAL PROBLEM: WATERSHED MANAGEMENT

The following characteristics will serve initially to explain our view of an environmental management problem:

- there is need for joint coordinated action for collective survival,
- over a relatively long term period,
- with equity of benefit a requirement.

These requirements are, of course, ambiguous. What exactly is "collective survival"? How long is "relatively long"? What exactly is "equity"? Should equity be weighted by population? Should past deficits be compensated or "wrongdoings" "punished"? We note that the achievement of equity typically requires inequality of restraint. Inequality, either across the society or through time, requires for its resolution sacrifice by some. How is that sacrifice to be achieved given that the agents are self-seeking?

As a major example of environmental resource management we consider *integrated watershed management*. This is the task of organising the activities and requirements in a river basin to achieve multiple and conflicting goals [Westervelt, 2000]. Stakeholder cooperation is essential. Typically there are conflicting requirements to be balanced of water supply (for domestic, agricultural, industrial uses), pollution control, fisheries management, flood control, hydropower production, navigation and wetlands management, recreation provision, and more.

Always there will be many stakeholders associated with different activities in the basin, all with their own objectives and agendas. Conflicts of interest are inevitable. A good example is the Fraser River basin in British Columbia [Healey, 1999; Doran, 2001].

3. INTERVENTION AND PERSUASION

Integrated watershed management, and similar ecosystem management problems, typically involve *intervention*. That is, some person, some group or some organisation, undertakes the task of intervening in the ecosystem (*including its human components*) in order to bring about desirable change, often using the notion of a search for *sustainability*. The intervener may be, for example, a branch of the UN, an NGO, an academic research team or even a lone doctoral student. The practice of intervention is so much a part of the ecosystem management task that, in our view, it is unrealistic to ignore it for modelling purposes. The intervention history of the Fraser River basin, just mentioned, is a revealing example of what issues can arise in intervention, and what can go wrong [Dorcey, 1997; Marshall, 1998].

The intervener is rarely if ever in a position of absolute control so that persuasion is required. It is evident that there can be a range of *intervention strategies*. A search of the literature [see Doran, 2001], and common sense, suggests at least the following possible intervention strategies:

- Apply "leverage" to some or all stakeholders to "force" cooperative behaviour by way of threats, punishments or rewards.

- Expose (more of) the reality of the situation to some or all stakeholders in the belief that the desired individual and collective behaviour will follow.

- Direct individual stakeholders' attention to the benefits to them of the desired action, avoiding mention of the consequent disadvantages. This exploits the cognitive limitations of stakeholders.

- Misrepresent the situation to some or all stakeholders but so that they act as desired out of self-interest.

Although aspects of these strategies may seem Machiavellian, they are surely commonplace in practical politics and social action, and as such are all worthy of study. Thus we would like, *with minimal prior assumptions*, to create a theory of social intervention strategies, for the particular context of environmental management, within which empirical observations such as these may subsequently be located. We therefore aim to model *all* of the intervention process on a computer in order to explore possible intervention strategies with the minimum of habitual and cultural pre-conceptions.

3.1 The Problem Restated

We now formulate the environmental resource management problem a little more precisely. We assume that exploitation of natural resources requires:

- distributed action coordinated in space and time

Furthermore actors (individual or organisational) must show restraint if they are to achieve, as we shall require:

- collective long-term survival (i.e. *sustainability*)
- the *protection* of specified environmental components
- some kind of *equity* between actors

The central difficulty is that human beings tend to be individually, collectively and organisationally "greedy" and with bounded rationality. In particular, we tend to think short-term. Any potentially informative model must capture these characteristics. Compare "common pool resource" (CPR) problems of which this formulation may be seen as a generalisation [Hardin, 1968; Ostrom, 1995].

3.2 A Research Plan

For clarity and focus, we foreground the following five-stage computer-based research plan:

1. Formulate a typical environmental system model (an ENVSYS) in mathematical and/or computational terms. Although abstract, the ENVSYS must reward distributed coordination and embody the sustainability, equity and protection problems identified above. Examine its long-term dynamics.

2. Automatically generate (on a computer) a sample of MAS connected to the ENVSYS. They should be neither incoherent nor successfully achieving sustainability, protection and equity over the chosen time span, that is, the generated MAS should function but *fail* to solve the problems.

3. Try to interpret the generated sample MAS in first abstract then human/social terms. This will probably include recognition of different types of MAS.

4. Search the space of all possible exogenous interventions to find those that are most successful for MAS of each type, where success refers to a high degree of maintenance of harvest, without depletion of protected environmental components, and with equal distribution of harvest over agents.

5. Interpret the interventions found in both abstract and human/social terms

Throughout the execution of such a research plan it would be essential not to confuse the computational and real-world domains. The central questions are whether effective intervention strategies can be identified in the computational domain, and then whether or not these identified intervention strategies have relevance to the real world domain.

4. A PRECISE FRAMEWORK

To proceed we need a precise and programmable specification of a MAS+ENVSYS, and of possible interventions upon it, that is sufficiently realistic for conclusions drawn from it to be of some practical value. The following framework specifies a suitable *class* of models rather than one specific model. We envisage use of a process of intelligently designed, heuristic and efficient "generate and test" to obtain specific models with the requisite properties.

4.1 ENVSYS

An environmental system, or ENVSYS, is to be structured as a set of Boolean, integer or real-valued variables inter-related by recurrence relations of the general form

$$x_n(t+1) = f(x_1(t) \ldots\ldots x_q(t))$$

where t refers to time and the subscripts index variables.

It is *not* intended that the ENVSYS be a model of a particular real-world environmental system. Rather the recurrence relations, together with the "actions" available to the agents (see later) and the agents' "localities" (see later), are chosen to provide the required resource management problem characteristics, that is, the need for distributed and coordinated harvesting together with difficulty in achieving sustainability, protection and equity. Distributed and coordinated harvesting is a matter of a specified pattern of actions upon a particular set of variables (actions and variables distributed in time as well as over localities) having a disproportionate and "beneficial" impact upon key harvestable variables. Motivating real-world instances range from large-scale irrigation systems and specialised artefact production to simple group cooperation activities such as ditch digging and tree felling. Problems of sustainability (and protection) may be posed by so choosing the ENVSYS relations that harvesting beyond a certain amount results in the harvestable (or protected) variables being driven beyond acceptable limits or permanently set to zero. Equity is naturally expressed as the requirement that all agents harvest to roughly the same degree.

The ENVSYS may be interpreted in many ways. For example, the recurrence relations may be read as a classical systems dynamics model [see Westervelt, 2000]. Alternatively, the ENVSYS may be read more in the tradition of "Artificial Life" studies with a spatial interpretation that has agents moving and harvesting localised resources on a plane [e.g. Epstein and Axtell, 1996].

4.2 MAS Agents

Each agent is structured as a set of tokens, the contents of its working memory (WM), together with condition-action rules that execute upon and manipulate the working memory and which observe and manipulate the agent's external context.

Each *token* is EITHER a simple token: a (bounded) string of letters, possibly prefixed by not (the negation character) OR a variable-value token: a pair consisting of a (bounded) string of letters, and a value.

Each *rule* is a pair consisting of a (bounded) set of tokens and a (bounded) set of actions, where an action is of one of five types:

- *Harvest* -- deplete a specified ENVSYS variable by a specified amount.
- *Set* -- set a specified ENVSYS variable to a specified value.
- *Read* -- read the value of a specified ENVSYS variable and deposit a corresponding variable value token in the WM.
- *Deposit* -- deposit a specified token in (own) WM.
- *Send* -- deposit a specified token in WM of another specified agent.

Each agent has its own locality, which is fixed in time, meaning that each agent can set, read and harvest a specified subset of the variables - its "local" variables. *Agents must harvest at a minimum total rate or they are deleted.*

4.3 Agent Processing

Agent processing consists of three main operations: rule firing, token reconciliation and rule set reconciliation.

Rule firing involves executing the actions of a randomly selected rule from the set of rules whose LHS match in the current WM. A match requires that every LHS token occurs in the WM.

Two tokens contradict if they differ only in the negation character. It is assumed that the initial contents of the WM are contradiction free. If a token is introduced (by an internal or external rule firing or an intervention) that contradicts an existing token, then the pre-existing token is deleted from the WM. This conflict removal procedure is very simplistic and certainly not, of course, logically complete in any technical sense.

Two rules contradict if their conditions are identical but their actions differ. It is assumed that the initial rules set is contradiction free. If a rule is introduced into the rule set (by intervention) that contradicts an existing rule, then the pre-existing rule is deleted. Again, this conflict removal procedure is not logically complete.

4.4 Intervention

An *intervention element* is the exogenous deposition of one token or one rule into a particular agent's working memory at time t. An *intervention* is a set of N intervention elements. The impact of an intervention element is determined by the reconciliation procedure.

4.5 Top Level Processing

After a MAS+ENVSYS has been initialised and the clock set to zero (t=0), the following sequence of operations is repeated until a time limit:

1. Advance the clock (t)
2. Activate each agent once in a varying random order
3. Pass any inter-agent messages
4. Apply any interventions at this time
5. Reconcile each agent's tokens and rules
6. Update the ENVSYS

This framework is computationally quite straightforward. In some respects, for example token and rule reconciliation, it could hardly be simpler and more arbitrary. Nevertheless, fleshed out with specific agent rules and localities, and with initial token sets for the agents' working memories, it is clearly programmable (in, for example, C++) and particular instances of it ("models") can therefore be run and tested.

5. MAJOR CHARACTERISTICS OF THE FRAMEWORK MODELS

We now consider the characteristics of these models, and the technical difficulties that any attempt to use them will encounter. Although it is not easy to anticipate in any detail what specific types of dynamics will occur within an agent or a MAS specified in these basic terms, some aspects of their behaviour are foreseeable.

5.1 Properties and Problems

It is important to appreciate that complex cognitive processes, for example the use of internal representations, goal setting, plan formation and execution, and learning, are potentially emergent in these agent's working memory dynamics even though these agents are "merely" rule based. This follows from the fact that the contents of an agent's working memory both determine the firing of, and are modifiable by, the rules. That does not mean, of course, that agents with cognitive processes are easily generated nor, less obviously, that it is easy to recognise them when they are. Indeed, just how cognitive processes can be recognised in practice in such a context is an interesting and far from trivial question.

The emergent behaviour of any particular model that meets the specified requirements is primarily determined by *the rule sets and initial token sets within the agents*. These rule and token sets must be such that the MAS, without intervention, has the specified properties with respect to the ENVSYS, notably that it does successfully "harvest" resources, but not so that harvesting is sustainable, equitable and protective. But the probability that an arbitrary or randomly generated set of agents (rule sets and token sets) for a MAS will function in this way, or even function coherently, is very small indeed. There is therefore a significant combinatorial problem merely to find functioning and effective MAS. As suggested earlier, some form of heuristic "hill-climbing" algorithm or evolutionary algorithm over the space of rule and token sets could be used. Just how complex are the effective MAS that could be found in this way is an open research question. Of course, one could set out explicitly to *design* and implement an effective MAS (using, for example, BDI or neural network architectures) but this would be to pre-determine exactly what we wish to discover. It would also encounter the standard difficulty that our ability to design the needed artificial intelligence capabilities is limited. A possibly effective compromise would be to design some basic structures and capabilities into the agents, perhaps sufficient for their minimal survival by purely uncoordinated action in the ENVSYS, but to leave the rest to some form of heuristic search over the space of rule and token sets.

Once generated, models meeting the requirements may or may not display collections of agents that may reasonably be labelled "organisations". They may or may not display centralised decision-making and/or collective planning. Agents (and agent organisations) will typically be heterogeneous, perhaps in a patterned way and, as just suggested, may or may not incorporate cognitive processes. Unfortunately all discovered MAS are likely to be "noisy" in the sense that their rules and working memory contents will often include much that is inessential to their required functioning. It is this property of ubiquitous "noise" which makes the problem of interpreting the internal processes of an effective MAS in terms of human cognitive and social characteristics so potentially difficult and interesting. Very little research has been addressed to this problem.

5.2 Interventions and Intervention Strategies

Recall that the purpose of generating MAS that can successfully interact with the ENVSYS (in the short term) is first to discover what form such MAS can take (rather than prejudge that issue) and then secondly to discover effective exogenous interventions. It may well turn out that we discover effective *patterns* of intervention (*intervention strategy)*. Note that an effective intervention strategy may directly prompt a successful environmental action, or it may bring about that action by first prompting the establishment of some suitable intermediate *social structure*.

For a given model (MAS+ENVSYS), optimal exogenous interventions can be defined and (in principle) determined without addressing the issue of an *intervener's* possibly partial knowledge of the model. However, this further and daunting issue cannot be avoided should we require a *decision procedure* that gives an effective intervention as a function of the intervener's knowledge of the model.

6. DISCUSSION

There are important advantages to this type of research programme, using this type of model. Such models are relatively objective, precisely because they are built of elemental structures, with their more complex behaviour emergent. They therefore help us to escape from preconceptions, including cultural and political preconceptions. They also help us to bypass pragmatic and habitual but nevertheless inessential and damaging intellectual barriers between particular schools of thought and between particular scientific disciplines such as psychology, computer science, artificial intelligence, cognitive science, and various of the social sciences. Furthermore, it is worth noting that Keohane and Ostrom [1995] have demonstrated the close relationship between cooperation for the solution of environmental problems, and more general international cooperation. Thus the potential research payoff is not limited to environmental contexts.

However, it may be argued that a study of this type can have very little practical value since only the simplest models can be found however sophisticated the heuristic search procedure deployed, and these models will therefore be unrepresentative and misleading. But this concern is unduly pessimistic. The models we suggest focus upon abstract essentials rather than potentially overwhelming detail. The success of heuristic search techniques in finding solutions to large combinatorial problems is demonstrated and well known. To assume that such methods will be useless in this context is unjustified. Furthermore, the structure of the search problem, involving specific and well-defined requirements that models must meet, means that the search is through a space that is in fact quite tightly constrained. Coupled with ever increasing available computer power, interesting discoveries are quite possible.

7. CONCLUSIONS

We have suggested how environmental management intervention strategies can be discovered and classified in the abstract by automatically generating and exploring a repertoire of relevant agent-based models. The objective is to match discovered abstract strategies to those in actual "everyday" use, and vice-versa, in an insightful and practical way. This includes intervention strategies that use certain types of social structure (or the lack of them) as "stepping stones". But there are major technical problems to be overcome of two kinds: exactly how to generate specific models of sufficient complexity to be representative and informative, and how to interpret models once generated. Thus although the potential payoff is too substantial to be ignored, the proposed research programme is a long-term and challenging one.

8. ACKNOWLEDGEMENTS

I am most grateful for the helpful suggestions of two anonymous referees concerning both content and presentation of this paper.

9. REFERENCES

Doran, J. E., From Computer Simulation to Artificial Societies, *Transactions of the Society for Computer Simulation International*, 14, 69-78, 1997.

Doran, J. E., Intervening to Achieve Co-operative Ecosystem Management: Towards an Agent Based Model", *Journal of Artificial Societies and Social Simulation*, 4(2), 2001, ON LINE: http://www.soc.surrey.ac.uk/JASSS/4/2/4.html .

Doran, J. E., S. Franklin, N. R. Jennings and T. J. Norman, On Cooperation in Multi-Agent Systems, *Knowledge Engineering Review*, 12(3), 309-314, 1997

Dorcey, A. H. J., Collaborating Towards Sustainability Together: The Fraser Basin Management Board and Program, in *Practising Sustainable Water Management: Canadian and International Experiences* (D. Shrubsole and B. Mitchell, eds.), Canadian Water Resources Association, 1997. http://www.interchg.ubc.ca/dorcey/chcwra/fccwra.html

Epstein, J. M. and R. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up*, The Brookings Institution Press, Washington, D.C., The MIT Press Cambridge, Mass., 1996.

Hardin, G., The Tragedy of the Commons, *Science*, 162, 1243-1248, 1968.

Healey, M., ed., *Seeking Sustainability in the Lower Fraser Basin: Issues and Choices*, Institute for Resources and the Environment, Westwater Research Centre, University of British Columbia, Vancouver, 1999.

Keohane, R. O. and E. Ostrom, eds., *Local Commons and Global Interdependence: Heterogeneity and Cooperation in Two Domains*, Sage Publications, London, 1995.

Marshall, D., Watershed Management in British Columbia: The Fraser Basin Experience, *Environments*, 25(2-3), 64-79, 1998.

Moss, S. and P. Davidsson, eds., Multi-Agent-Based Simulation *Proceedings of the Second International Workshop on Multi-Agent Based Simulation (MABS 2000)*, Boston, MA, USA, 8-9th July, 2000, Springer, Berlin, 2001.

Ostrom, E., Constituting Social Capital and Collective Action, *Local Commons and Global Interdependence: Heterogeneity and Cooperation in Two Domains* (R. O. Keohane and E. Ostrom, eds.), Sage Publications, London, 125-160, 1995.

Russell, S. and P. Norvig, *Artificial Intelligence: a Modern Approach*, Prentice-Hall, 1995.

Weiss, G., ed., *Multiagent Systems*, The MIT Press, Cambridge, Massachusetts, and London, England, 1999.

Westervelt, J., *Simulation Modeling for Watershed Management*, Springer-Verlag, Telos, 2000.