Simulating Societies using Distributed AI

Jim Doran
University of Essex
Colchester, CO4 3SQ, UK

ABSTRACT: This paper discusses the prospects for using Distributed AI techniques to support the computer simulation of societies. Newly developed ideas and techniques are reviewed, some relevant projects are briefly described, and some potential pitfalls are noted.

1.0 Introduction

Computer based simulation of societies is again a growing research area (Gilbert and Doran, 1994; Gilbert and Conte, 1995), with the use of distributed AI techniques a particularly promising and exciting recent development. However, given the convergence of research traditions involved, it is not easy to achieve a clear overview of the potential pitfalls and prospects. In what follows I shall attempt such an overview, making brief illustrative reference to ongoing work at the University of Essex.

2.0 Modelling & Simulation

What is modelling? In essence, an investigator sets out to use one entity, the model, as a tool for the study of another, the target system. Careful manipulation of the model provides information about it, which is then transferred to the target system, hopefully in a reliable way. Computers can support the construction of the model and its manipulation and the recording of the model’s behavior. In this account I shall not distinguish computer modelling from computer simulation -- but see Doran & Gilbert (1994).

There are at least three types of theory at issue in a modelling study: the theory of the target domain (here some branch of social science), the theory of the type or types of models used (quite likely a branch of mathematics, computer science or artificial intelligence), and the theory of the modelling process itself. Many modelling exercises contribute to
more than one type of theory, especially when, as is the case here, both the theory of the model and of the domain is incomplete.

2.1 The Space of Models and the Relationships within it

There are many different types of mathematical or computational system capable of being used as a model and many variants of each type. Some prominent candidates are sets of differential equations, cellular automata and Distributed AI style models (see below). How should one choose between them on any particular occasion? Unfortunately we have rather little understanding of the relationship of these models one to another, although it is clear that some models may be regarded as elaborations of others (e.g. Zeigler, 1990). Hence, one frequently heard slogan is that:

*Models should be made as simple as possible but no simpler!*

carrying the implication that a model which is sufficiently rich to answer the questions to be asked of it need not and should not be further elaborated. Realism for its own sake is misguided and will merely confuse and generate computational (over)load.

2.2 Will the REAL model please stand up!

A major hazard in modelling work (for both social scientists and AI scientists) is for the investigator to fail to separate a model’s actual symbolic structure and dynamic properties from the modelling interpretations that are being associated with it.

It is all too easy to read a sophisticated interpretation into a relatively simple formal model, and then to believe that conclusions based upon the interpretation are independently supported, when in fact the model merely illustrates the investigator’s preconceptions. It should always be possible to state the properties of the model (possibly newly discovered properties) independently of the target system at which it is aimed.

3.0 Distributed AI

Distributed AI (DAI) is the study of what happens when a set of “intelligent” computational entities (“agents” in a “multiple agent system” or “MAS”) are allowed to interact and possibly intercommunicate (O’Hare and Jennings, 1995). By “intelligent” is meant that the agents embody computational processes which may reasonably be described as forms of perception, planning, learning, decision making and so on.

The objectives of DAI are:

- to establish the (precise and formal) properties of MAS
- to build useful MAS for applications such as air traffic control, cooperative engineering, and distributed sensing.
- to use MAS as models of naturally occurring multiple agent systems.
It is, of course, the last of these alternatives that is primarily at issue here.

I shall introduce the content of DAI by sketching some typical agent designs, and some typical ways in which sets of agents may be organized. But it must be emphasized that building “cognitive” abilities into computational agents (which is the mission of AI research) is no easy matter. Much progress has been made and prospects are good (see, for example, Russell and Norvig, 1995; Wooldridge and Jennings, 1995) but nevertheless the cognitive abilities of existing computational agents are greatly limited.

4.0 Agent Designs

Agents may be characterized as mechanisms embedded in an environment that select and perform actions in the environment in the light of their current input from the environment (perception) and their current state of memory/belief/knowledge.

A variety of different agent “architectures” have been designed in recent years and their properties partially explored. I shall briefly describe three of them. The list is not exhaustive -- it does not include, for example, neural network based agents nor ‘fuzzy logic’ controllers.

4.1 Basic reactive agent

One type of simple agent architecture, based on the notion of a “production system” (see Ishida, 1994) consists of the following main components:

- a ‘working memory’ (comprising a changing set of tokens derived from “perception”)
- a set of rules of the form IF <condition> THEN <action>. Such rules are often called “production rules” or “condition-action” rules. The condition part of the rule typically specifies a required conjunction of tokens (the specification may involve variables) and the action part a specification of an executable procedure.
- a mechanism that repeatedly identifies a rule whose condition part matches the contents of the working memory, and then executes that rules’ action procedure -- thereby causing the agent to perform or more actions in its environment.

Thus the essentially functioning of the agent is that the results of perception are deposited as tokens in the working memory, and rules are ‘fired’-- with the result that actions are performed by the agent in its environment.

The heart of such an agent is thus the set of condition-action rules which it embodies. The effectiveness or ‘intelligence’ of the agent, if any, is entirely a matter of the care and success with which the rules have been devised, and the knowledge of the environment which they embody.
4.2 More complex agent designs

The basic agent structure just described can be extended in many ways. Two of the more important are:

- maintain a semi-permanent body of knowledge in the working memory and include rules which update the contents of the working memory -- so that the agent is, in effect, able to maintain and update some sort of representation of its environment.

- design some of the condition-action rules so that they post goals and construct plans of action in the working memory and then execute a plan selected from those constructed. The agent then becomes, in a certain sense, *deliberative*.

4.3 The Fashionably Layered Agent

Currently there is considerable interest in layered agent design -- essentially several agents of relatively simple design are coupled together as layers. The layers might be a low level basic reactive layer, a deliberative layer and finally a high level layer intended to handle cooperation with other agents.

This is the agent design espoused in, for example, the well known InteRRaP project (Muller at al, 1995). Its attractiveness is that it offers a modular design, with clear relationships between modules (the layers), rather than a hazardous attempt to combine multifunctionality in one integrated system.

4.4 Agents with Beliefs and Emotions

The agents designs just described easily become structurally complex but are, at the same time, prey to a dangerously simple assumption -- that rationality and knowledge must be the corner stones of any “intelligent” agent in any context. There are at least two ways in which such an assumption may be mistaken: as regards belief rather than knowledge, and as regards the role of emotions (or “affect”):

- in most environments and in multi-agent setting agents must deal in beliefs rather than knowledge. Given the unreliability and limited nature of perception (and communication) an agent will typically have to rely on beliefs about the environment and other agents which are partial and often erroneous -- *misbelief* is the norm. This has major implications for agents design.

- even more controversial is the issue of the role, if any, to be played by “emotions” in computational agents. In human society it is at least plausible that emotions play a significant role in determining macro as well as micro behavior. This suggests that affect systems must be part of our models, which is, in principle, quite possible (arguments along the lines: “but computers cannot have emotions” miss the point that we are *modelling*). Agents can be given “emotional” variables which impact cognitive processes and which are to a lesser or greater extent communicable -- see below, section 7.2, and,
for example, Doran and Clarke, 1993; Ortony et al 1988. Whether an intelligent agent will perform better, in some sense, when equipped with emotions is a separate question, yet to be resolved.

4.5 Other Perspectives on Agents

The view of agents I have just given is (emotions and beliefs aside) probably the dominant one in Distributed AI. However even within DAI there are alternatives: for example, the BDI approach (Beliefs, Desires, Intentions) puts much more emphasis on formalization as a preliminary to (hoped for) agent design. And outside Distributed AI elements of the mainstream computer modelling community have also begun to consider “intelligent” agents e.g. Zeigler (1990) and see Uhrmacher(these proceedings).

5.0 Sets of Intelligent Agents

Distributed AI is concerned, of course, not with agents in isolation but with sets of agents. There are two extremes:

- communities of autonomous agents which share a (possibly simulated) environment and which compete within it. This may lead, for example, to the study of inter-agent negotiation protocols, and also of ways in which essentially competitive agents may believe it mutually beneficial to cooperate and act accordingly. Some Artificial Life work is of this type.

- cooperative and tightly organized agents, for example in a hierarchy. Such organization raises issues of task delegation, responsibilities, roles, and the degree of awareness the individual has of the whole. Classic techniques of DAI research include the Contract Net Protocol, Partial Global Planning, and Generalized Partial Global Planning (see O’Hare and Jennings, 1995).

There are a range of possibilities between these two extremes. Especially relevant to the study of human societies are computational scenarios where many partially organized communities exist and compete in a shared environment.

6.0 Using DAI to Simulate Societies

It seems reasonable to argue that a model of human society and its dynamics should embody distinctively human cognitive attributes. Otherwise, human society is modelled as if it were, say, an insect society, and this is likely to be effective only in special circumstances. We need AI style agents in simulations so that the processes of cognition and their implications (including side-effects) are explicit and not simplistically assumed.

DAI software testbeds support experimentation with multiple agent systems (Decker, 1995). They can and should be used in experimental mode to discover new relationships between structure and behavior in social systems. I shall return to this key point later.
7.0 Illustrative examples

The following work is ongoing at the University of Essex. My aim in briefly describing it is to give the flavor of what is currently possible.

7.1 Tsembagan decision making

Rappaport, in his classic anthropological study *Pigs for the Ancestors*, describes decision making amongst the Tsembagan people of New Guinea. Faced with a major decision to be made, for example whether or not to begin the building of a new guest house, men discuss in small fluctuating groups with seemingly random movement between the groups. The groups are small -- 3, 4 or 5 persons -- and may or may not be localized in one area. Finally someone takes a lead, presumable feeling that a near consensus has been achieved -- and may or may not be followed (Rappaport, 1984, pages 30-31).

Grew (1995) has created a simulation of this scenario, in two dimensions, involving simple agents who can hold one of a small set of alternative “opinions” about a range of “topics”, and who change opinions and move between groups in accordance with the distribution of opinions within their current group and with probability specifications. For example, an agent tends to take on the majority view of its group about a particular topic, and to move to another group if either there is unanimity in the group or its opinion is strongly in the minority. Opinions about different topics might or might not be consistent. The final decision to “take a lead” is made by an agent if recently it has encountered only other agents sharing its own (consistent) set of opinions.

In systematic experimentation it has been possible to investigate the impact of such independent variables as:

- group size,
- the shape of the total meeting area,
- the size of the range of opinions that might be held,
- and the extent to which the available set of opinions is inconsistent

on the time taken until termination is reached (i.e. an agent took a lead) and the degree of agreement with the lead opinion at termination.

An unexpected insight obtained from Grew’s work is that the agents’ collective decision process may be seen as a special case of distributed constraint satisfaction problem solving (e.g. Tsang, 1994) wherein many agents are collectively seeking to find a solution to a problem posed as variables and constraints, each agent having, however, a view of only a small part of the problem structure. This opens up the prospect of returning to the anthropology, and seeking more evidence of people solving this type of problem in this manner.
7.2 Belief and Affect in Hierarchical Organizations

A branch of Organization Theory known as computational organization theory has recently emerged (Carley and Prietula, 1994). There is much ongoing computer-based work, often aimed at linking the design of an organization to the characteristics of its market environment. The typical assumption is, however, that organizations are composed of agents working rationally with relatively reliable information. There is surely some interest in relaxing this assumption.

At Essex, Hammond (1995) has used a computer simulation to study the properties of a hierarchy of agents, where individual agents in the hierarchy are also randomly located in two dimensional space. “Tasks” are submitted to the agent at the apex of the hierarchy, are delegated down it from one agent to another, and then performed at the lowest level. Agents (a) have beliefs about the workrates of those spatially close to them, and (b) have, in a certain limited sense, ‘feelings’ about their situation. What this means is that the investigator can, for example, arrange that any agent believing that those around it have lower workrates (both in performing and delegating tasks) than itself, becomes “angry” and, in consequence may, as determined by the investigator in any particular experimental trial, work less hard, harder, or more erratically. Further, the beliefs that agents have are subject to error and are communicated from one agent to another. Hence the possibility arises that all the agents lower in the hierarchy, say, or in a certain spatial area, say, believe rightly or wrongly that the senior management as a group have very low workrates and react accordingly in “emotional” terms. The scientific interest, of course, is in the connection between micro-level specifications of agents and their reaction to their circumstances, and the macro-level behavior of the hierarchy as a whole.

Initial experiments and their results have suggested that a degree of misbelief (in this case error) by agents about the workrates of those around them improves the performance of the organization as a whole. At first sight this is a surprising result and needs confirmation. It is not, however, totally implausible. It may be, for example, that a degree of error in agents’ awareness of what is happening around them is sufficient to block negative feedback loops that would otherwise take effect. But this is a speculation. It is not always easy to work out what is happening in multi-agent simulations, and work is continuing -- see Rosales (forthcoming)

7.3 The Emergence of Human Social Complexity

7.3.1 The EOS Project

The EOS project (Doran and Palmer, 1995) is an investigation, using DAI based modelling, of the origins of human social complexity. Initially we have set out to create and explore a computational interpretation of the model informally specified by Paul Mellars, a Cambridge archaeologist, for the growth of social complexity in the Upper Palaeolithic period (around the time of the last glacial maximum) in Southwestern France. Mellars’ model relates changing features of the natural environment to the emergence within the human population, for perhaps the first time, of centralized decision making, hierarchy
and related social phenomena. It emphasizes *population concentration* as the primary causal factor.

The EOS project makes heavy use of DAI concepts and techniques, and uses a dedicated software testbed written in Prolog which supports a community of agents in a simulated environment with “resources” which may be “harvested”. Following Mellars, we have focussed on the need for individuals (a) to *cooperate in resource acquisition*, and (b) to *conceptualize effectively their physical and social environment*. Agents are structured as production systems and are deliberative. Amongst other things the rules maintain agent-specific social models, that is, simple dynamic representations of other agents and of their social arrangement, and construct multi-agent resource acquisition plans.

Our results tend to support Mellars’ views. We have also compared and contrasted Mellars model, in computational terms, with the seemingly very different and also informal model of the palaeolithic archaeologist Clive Gamble. The two models turn out to be less in conflict than at first appears.

### 7.3.2 Kinship

Issues of kinship are not addressed in the Mellars model itself, but many others have suggested that the origins of social structure partly lie in the extension of family relationships and of the *concepts* associated with family relationships, to wider communities (e.g. Read, 1995). Mayers (forthcoming) is pursuing this possibility in the context of the EOS work, with particular regard to the study by Todd (1985) of the correlation between categories of family structure and categories of social system.

### 7.3.3 The Role of Collective Belief and Affect in Emergence

The issues of cooperative resource acquisition and population concentration are surely not the whole story (if part of the story at all!) of the emergence of human social complexity. There is also the issue of what collective belief system may be associated with such emergence phenomena, and what affect it embeds (compare Alexander, 1989; Hammond, 1990; and Renfrew, 1994).

The artificial societies created in the EOS testbed typically display the phenomenon of collective misbelief. The agents communicate and share beliefs (for example about the locations of resources, or the nature of other agents), but these beliefs are often false in the sense that they do not correspond to the reality of the simulated environment.

Experiments arising from the EOS work (Doran, 1994) are addressing the significance of collective misbelief in communities of computational agents and how such misbelief can be to the community’s benefit. It has been demonstrated how groups may be formed around pseudo-agents -- that is around entities in the environment which the community collectively and *wrongly* believes to be an agent. In certain circumstances such misbelief is beneficial to the members of the “cult” and to the wider society of which they are a part (Doran, in preparation).
It seems reasonable to suggest that an understanding of collective misbelief from this computational perspective can strengthen existing theories of ideology.

8.0 Prospects and Problems

The Essex work just sketched, although incomplete, suggests that a very wide range of human social phenomena (including those centered around social belief and affect systems) may be studied using Distributed AI based modelling. The prospects look exciting, but we are a long way from having understood the best way to proceed and how best to avoid pitfalls.

It is important to keep in mind that social modelling need not be merely a process of checking, or filling in the details of, pre-existing social theory. Assuming a commitment to the experimental/modelling paradigm -- which, alas, is contrary to the preferences of many theoreticians in both DAI and the social sciences -- simulation gives us the opportunity to discover new theory, to find what is the case in a more experimental mode than a subjective assessment of empirical data. From the assumptions built into the model, emerge patterns of model behavior we may well have not anticipated. Not only that, by objectifying the process of deriving conclusions from assumptions, we are enabled to escape in some small measure from our own misbeliefs.

Major problems in practice are the complexity of the software needed to run computer-based experiments of this type, the time such experiments take, and the sheer multiplicity of data which they yield. These may not seem deep problems, but they are. In particular, discovering what really happened in one's multiple agent simulation is both essential and hard. It is beginning to seem that the fundamental problem at the abstract level is how best to identify and characterize core regularities in complex multivariate (and, of course, partially non-numerical) multiple agent systems, and so to propose new social theory.

9.0 Acknowledgments

The EOS project has been supported by the Joint Council Initiative in Cognitive Science/HCI (Grant no 8930879) and the EC CINECA ICARUS programme.

10.0 References


Mayers S, Forthcoming Masters dissertation, Department of Computer Science, University of Essex, Colchester, UK.


Rosales E, Forthcoming Masters dissertation, Department of Computer Science, University of Essex, Colchester UK


