The Archaeology of Artificial Societies

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

Abstract

Can archaeologists help software engineers unravel what has been happening in an artificial society of intelligent agents? We discuss the methods that archaeologists regularly use and how they relate to the properties of an artificial society and the problems faced in recovering its history. As part of the discussion, an abstract model of a typical artificial society is presented, the structure of the process of interpreting evidence is analysed, and the particular macro-social phenomenon of socio-cultural collapse is considered.

1 Introduction

Archaeologists ask: "What did those guys (humans) DO the last few millennia?"

Software Engineers (will) ask: "What did those guys (agents) DO the last few hours/days?"

In this short paper I want to explore an idea that might at first sight seem a little bizarre: that archaeologists may have something to teach those who have to debug, or just understand, what has been happening lately in some specific *artificial society of "intelligent" agents*. It may be timely and useful to study how archaeologists recover past human social processes, in the belief that this will help future software engineers recover the past histories of the artificial societies that they are tasked to manage. Obviously, this task of recovery will be especially important if the societies in question have been malfunctioning. And, just because we *are* addressing societies, we may assume that the process of recovery, often preparatory to taking some kind of remedial action, will typically be pitched primarily at the social rather than the individual or the code level.

It perhaps needs emphasising that recovery of history in this context is indeed going to be a problem. Even if, implausibly, we assume that comprehensive tracing/logs exist in an artificial society -- but the supporting infrastructure would collapse? -- finding out what has been going on and why will *not* be easy. The problems are (a) the sheer magnitude and complexity of the raw activity logs that must be worked through, including traces reflecting the internal processing of agents, and (b) the restricted and costly access to the raw logs in practice. These problems exist whether or not the process of scanning activity logs is largely automated. Think of the "pile of printout" generated by even 10 minutes of a multi-agent simulation involving many non-trivial agents, if the tracing of the simulation and the agents within it is at all detailed!

2 What do Archaeologists do?

At the heart of archaeology is a set of techniques for working back from surviving evidence to the social processes, located in time and space, which gave rise to it. Crucial is the fact that some objects and types of material (especially stone, bone and pottery) can survive for very long periods of time in the ground. Archaeological excavation therefore enables the recovery of past activities, but the process of interpretation of archaeological data must allow for partial and differential survival. Further, archaeologists can only sample a small part of what is now in the ground, and frequently find themselves excavating a small faction of the whole site which they know to be there. This most notably occurs in the context of "rescue archaeology" when an archaeological site is about to be destroyed by some kind of building work. Thus the process of acquiring archaeological data is always time-consuming and subject to biases and problems.

Over the last hundred years and more, archaeologists have evolved a meticulous methodology involving systematic excavation, recording (including of stratigraphy in the ground), handling and restoration of artefacts, and documentation. A feature of archaeological method is that excavation and recording must proceed without too much prior assumption of what is important. The most seemingly insignificant fragment may prove highly informative.

Archaeologists first seek to establish what is in the ground, then work back to what existed in antiquity, and how and why it got there. The process of archaeological interpretation may be regarded as addressing three levels:

- the raw excavation data -- what is found and its context
- the micro (human) level of interpretation -- the human activities (e.g. cooking, burial, flint working, slaughtering) that the raw data reflect
- the macro (social) level of interpretation -- the rise, stabilisation and collapse of complex societies, the migration of a populations, and so on.

Some of the main archaeological techniques in standard use are:

- The recording and interpretation of stratification during excavation. This enables the original interrelationship in time and space of different deposits to be inferred.
- Comparisons between artefacts (for example, stone tools) leading to typologies, seriations and hence relative chronologies.
- Interpretations (e.g. as graves/hearths/hunting camps/kilns) based on common sense, and on documented examples of modern societies which are judged similar to (some of) those of antiquity.
- Absolute dating. Some materials surviving from antiquity may be dated with some precision. Two of the most important techniques are carbon-14 dating and dendro-chronology.

For detailed examples of these technique and their use, refer to any archaeological textbook (e.g. Doran and Hodson, 1975; Renfrew and Bahn, 1996). A discussion of a specific and typical piece of archaeological reasoning, and of how it might be automated, may be found in Doran (1970).

3 Artificial Societies

We may here take an *artificial society* to comprise located software agents, fixed and mobile, which are heterogeneous and which inter-relate. An artificial society is continuously active. Agents may be simple or complex, reactive or deliberative. If, as we may assume, the society is open, then agents enter and leave the society unpredictably. Agents may also self-clone or otherwise reproduce. There are potentially complex tasks to be performed, perhaps involving "documents" or other items, which are created and deleted, passed from one agent to another, exchanged, imported and exported, transformed and combined. Between the agents there may well be permanent and semi-permanent relationships and commitments, including dominance. At the social level (see Jennings and Campos, 1997), co-operative groups, organisations, markets and other social structures may either be designed into the society or may be *emergent*.

One widely discussed and plausible vision of the "post-PC" era of computing (e.g. CPHC, 2000) is of a world in which computing is distributed and networked throughout the environment in which we live, and is incorporated in the most mundane objects such as refrigerators, cars, door locks and even clothing. As these "things start to think" (in the current phrase), it seems that the post-PC vision is inevitably one of massive agent societies. It seems unrealistic to assume any kind of central control of such a society. Rather the infrastructure and the agents that it supports will be subject to diverse origin, ownership, and management, but with network wide standards and conventions to ensure coherence.

3.1 The Problem

What should happen if some kind of recent failure of the artificial society's functioning is detected, for example, if the total amount of useful activity has suddenly declined? To understand what has gone wrong, what has happened in the past must be understood.

What kinds of activity log will a "massive" agent society leave? It is safe to assume that the logs will be partial and heterogeneous, and that a software engineer will have access only to part even of the existing logs, which will be costly to obtain. We might reasonably assume that there is available:

- some evidence of agent location and movement
- some evidence of inter-agent communications and what they are about

but that as with humans:

- there is *little or no* record of what has been happening *inside* the agents.

3.2 An Abstract Model of an Artificial Society

In order to formulate the recovery problem more precisely, we may usefully consider the following *abstract model* of an artificial society and its associated recovery problem:

> There is a (very large) *network* (graph) whose nodes we call *sites*[1] and whose (bi-directional) links we call *channels*. At any moment many *agents* and *items* are located at sites on the network.
>
> Agents and items (and messages, below) have unique identifiers.

[1] This use of the word "site" is intended to echo the standard use of the term by archaeologists to denote any location of archaeological interest.

Each agent belongs to a certain *class* of agents (e.g. has a single, external owner). Many agents may belong to the same class.

Agents obey unobservable (by the software engineer) *internal decision rules*, which vary from one agent to another.

From time to time agents spontaneously appear at and disappear from sites. Thus the society is *open*.

Some agents are fixed (located permanently at one site). Some are able to move between sites (via channels).

Agents exchange messages of various types (via channels).

Items are of a number of types.

Agents perform the following actions upon items:

- create or delete an item (if agent and item at same site)
- pass an item from one agent to another via a channel (or two agents exchange items)
- import or export particular types of item from particular sites in the network
- transform an item from one type to another (if agent and item at same site)
- combine two or more items of appropriate types into a single item of a different type (if agent and items at same site)

There exist *item combination constraints* (known to the software engineer), which determine the outcome of the agent actions to combine items.

At each site there is maintained a (recent) history of (the identifiers of) incoming and outgoing agents, incoming and outgoing items and their types, and incoming and outgoing messages and their types. The content of messages is not recorded. At a moment of termination there is available to the software engineer a subset of these site histories.

The software engineer may assume that the intended *function* of the society is that each class of agent should create and export certain specific types of item.

The software engineer's task is to reconstruct as much as possible of the activity in the society, at a suitable level of abstraction, in the period up to a moment of termination.

3.3 Comment on the Abstract Model

It should be apparent from the model specification that the agents in the society are essentially tasked to create complex items by the appropriate use of *combination actions*. It follows that organised co-operation between the agents in the society will much enhance their effectiveness. Understanding of the society may therefore focus upon looking for stable patterns of co-operation and, perhaps, their "collapse" (see section 6.1). In turn these patterns of co-operation may well be reflected in patterns of messages recoverable from the site histories available to the software engineer.

For example, suppose that a group of agents is able to use a version of the well-known contract net protocol to perform a specific item compounding task, that is, many items of various types must be subjected to a set of combination actions so that a particular type of item is generated. One agent initiates the task and delegates item retrieval and item combination tasks to others and this process of delegation is repeated recursively. One may predict that a *distinctive pattern of messages* will often be recoverable from the assumed activity logs.

The nature of the items within the model is deliberately left open here. Item combination may loosely be compared with the construction of, say, an automobile. But items in artificial societies are in reality more likely to be text documents of one kind or another.

4 From Archaeology to Artificial Societies

Can we now map what archaeologists do onto what software engineers must do to unravel the history of an artificial society? We consider some important correspondences one by one.

4.1 Targeted and "Rescue" Excavation

As stated earlier, archaeologists are frequently obliged to excavate sites on an opportunistic basis, prior to the sites' destruction. For artificial societies this corresponds to the likely limited availability of site activity logs and to time constraints on the recovery problem, perhaps imposed by site log deletion procedures. The archaeological use of formal statistical sampling may carry over to the agent domain.

4.2 Excavation Technique

Archaeological excavation corresponds in artificial societies to the recovery of detailed activity logs from particular network sites. Although it seems likely that these records will be easier to assemble and interrelate than are the stratigraphic and other contexts that archaeologists must deal with, this may not always be the case, and general archaeological requirements of meticulous study without undue prior assumptions will apply.

4.3 Artefacts, and Relative and Absolute Chronology

The notion of an individual artefact has at least two possible analogues in artificial societies. Firstly, it may perhaps be associated with individual log entries. However, in artificial societies we can probably assume the availability of absolute dates/times for log entries as the norm. This implies that much of the archaeological concern with establishing chronology via such techniques as stratigraphic analysis, and artefact seriation based upon typology will not arise.

Alternatively, and perhaps more persuasively, an association may be made with the "items" included in the abstract model of an artificial society presented earlier. Then the construction of typologies and seriations over sets of artefacts might arise as a means to the understanding of the patterns of co-operative "work" upon items that a group of agents have evolved between them.

4.4 Micro-Interpretations

As indicated earlier, archaeologists regularly interpret raw excavation data in terms of basic human activities such as cooking, burial of the dead, hunting and so on. The corresponding activities in artificial societies seem to be such micro level agent interactions as communication, negotiation, delegation and argumentation. To recognise instances of such interactions algorithmically seems quite feasible.

4.5 Macro-Interpretations

Encompassing micro-level basic human activities there are macro-level social phenomena. Archaeologists address these where they can. For example, they examine the existence of different types of society at particular locations and times (for example, centralised and/or ranked), the relationships (for example, trading) that may exist between different types of society, and the processes that appear to have led to changes, for example, migration or social collapse.

Similar macro-level phenomena are to be expected in artificial societies. Large communities of co-operating agents are to be expected and these may be recognised, and the complexity or otherwise of their internal structure assessed. But there will be at least one important difference. A particular agent community may well have a single external “owner” which directly or indirectly sets its “top goals” (as allowed for in the abstract model of Section 3 – the notion of agent classes). External ownership of agent communities will presumably constrain the macro-dynamics of the society in ways yet to be understood.

Some important indicators that archaeologists use to recognise macro-structure, for example developed societies, are:

- the complexity and size of sites (and their length of existence)
- evidence of trade links -- for example, artefacts all or part of which are remotely sourced.
- evidence of ranking and specialisation (for example, in graves)
- indicators of "civilisation", for example, writing and monumental architecture.

All of these indicators, except perhaps the last, can be giving meaning in the artificial society context. For example, the complexity of a particular agent community is measured by the proportion of its agents which can be shown to be specialised to a certain type of item processing i.e. disproportionate use of particular combination actions.

5 The Process of Interpretation

There are clearly certain similarities between archaeological interpretation and the interpretation of activity logs from artificial societies. These may be summarised by saying that both involve a set of *interpretation rules*, which must be used to recognise certain *entities* (e.g. hut, burial, auction, negotiation, migration). From an artificial intelligence perspective, the combination of rules and ontology (a “conceptual repertoire”) may loosely be regarded as a *frame hierarchy* (sometimes called a *schema hierarchy*) in which each frame contains both a characterisation of its corresponding entity, and also procedures for that entity's recognition, and in which the relationship which structure the hierarchy is *a kind of*. Interpretation is then a process of heuristic instantiation of some the frames in the hierarchy. This concept of a frame hierarchy with attached rules has been explored and implemented in, for example, the classic expert system CENTAUR (Aikins, 1983) in which the concepts represented in the frame hierarchy were disease entities. A closely relevant archaeological example is the PALAMEDE system of Francfort (1990) which addresses the archaeology of proto-urban eastern civilisations in about the Third Millennium BC. PALAMEDE "simulates" archaeological interpretation to the point of the recovery of macro-social dynamics (see next section), specifically the evolution of urbanisation.

It must be stressed that although in different problem domains there is similarity in the form of the interpretative process and its reliance on a combination of interpretation rules and an ontology, the actual rules and the actual ontology will surely differ from domain to domain. Thus the rules by which an archaeologist recognises the existence of, say, a prehistoric hut from traces in the ground, are quite different from those we would (implicitly) use to recognise a hut in existence now, for example by looking at it and doing some visual processing. And, of course, archaeologists work with specialised concepts (e.g. a "Levallois point", a "horizon") which do not exist in the everyday repertoire at all. The implication is that *recovering the history of an artificial society will also require the development of a conceptual repertoire and associated interpretation rules, conditional on the types of activity logs available for study.*

6 Understanding Social Dynamics

In the preceding section I suggested that the process of moving from evidence to interpretation involved rules of interpretation and an ontology that includes processes. But at the macro-social level the ontology and its associated processes, that is, the social dynamics, are not well understood in either human or artificial societies. Thus even with good information about low level activities the macro-dynamics are potentially very hard to recover and understand. It is therefore nor surprising that archaeologists tend to be cautious at this level. In particular, they rarely speculate about social trajectories that might in principle be quite possible, but which happen not to have occurred in prehistory. This means that social theory from the perspective of the prehistoric archaeologist is tied quite closely to archaeological record as it exists.

6.1 Socio-Cultural Collapse

As an example, consider the important and much studied example of a prehistoric macro-social phenomenon is *socio-cultural collapse* (Renfrew, 1979), which occurs when an established and complex society relatively suddenly disappears from the archaeological record or, at least, becomes sharply diminished in its complexity. There are many instances of this phenomenon in the archaeological record, of which perhaps the best known is the collapse of the Mayan society in Central America towards the end of the First Millennium AD. Socio-cultural collapse is particularly relevant here because were it to occur in an artificial society, it might be expected greatly to diminish the society's useful activity.

Archaeologists have identified many possible processes of collapse, some purely internal, some including one or more external factors. These have included not only such obvious candidates as invasion, disease, and climate change, but also more subtle "domino" effects impacting population centres, and negative feedback loops within the actions of the ruling elite.

Experiments with the EOS multi-agent system (Doran and Palmer, 1995) have suggested that two broad categories of collapse are:

> *Change in the environment of the society, such that its structure ceases to be functional.*

> *In a society where agents reproduce and "die", failure of the society successfully and continuously to reconstruct itself.*

Much more research is needed into trajectories of collapse in artificial agent societies, with the emphasis placed on identifying categories of possible trajectory, rather than merely modelling observed real-world instances.

6.2 Emergent Social Complexity

Those tasked to engineer effective artificial societies probably need to understand a wider range of macro-social phenomena than do archaeologists. Unfortunately almost everything remains to be understood, including much about the origins of emergent social complexity. In this regard we may speculate that (i) agent reproduction is important, and that (ii) so is the ability of an agent to "sell its own labour". Reproduction enables collective evolution and the emergence of agents with co-operative characteristics that may not be immediately predictable. By "selling its own labour", we mean that an agent, *in what it judges to be its own "top-goal" interests*, agrees a deal in which it makes a semi-permanent commitment to another agent's goal set[2]. Such an agent ceases to be fully autonomous and becomes, in effect, merely the occupier of a role (in one interpretation of that term). For example, an information seeking agent may, if it has the "authority" and ability, rationally choose to occupy a "role" in this sense in return for an information "feed". Thus, we may speculate, roles emerge and hence organisations as composites of roles.

But do the benefits of enabling such processes in the agents of an artificial society outweigh the danger that emergent phenomena will deflect the society from its intended function? Recent discussions of agent-based software engineering methodology (e.g. Jennings, 1999; Wooldridge, Jennings and Kinny, forthcoming) are cautious on this point and tend to assume that agent system designers will wish to exclude potentially uncontrollable emergent phenomena. If this view prevails then the problem of history recovering may be kept relatively simple.

7 Conclusions and Future Work

It seems clear that it is productive to compare archaeological methodology with software engineering methodology for artificial societies, and that there are two directions for further more detailed work. The first direction is to explore in greater detail the comparison between the standard archaeological process of data interpretation and that needed to interpret activity logs of agent societies. Particular topics are (i) the development of more precise abstract models of artificial so-

[2] For steps in this direction, see the Generalised Partial Global Planning co-ordination mechanisms of Decker and Lesser (1998).

cieties together with definite algorithms able to perform the recovery task for them, and (ii) a more detailed and insightful study of the relationships between the ontologies and interpretation rules corresponding to the two cases. It may be, for example, that the assumption I have made here that there will be no record available in artificial societies of the *internal* processing of agents, nor of the actual *content* of messages, will need to be revised.

The second direction for future work is to further investigate similarities in the macro-level dynamics of human and of artificial societies. Greater understanding seems possible and likely to impact both theoretical archaeology and the design of artificial societies and, indeed, to contribute to the development of general social science.

Acknowledgement

I am grateful to Omer F. Rana for encouraging me to pursue this line of thought.

References

J. S. Aikins. Prototypical Knowledge for Expert Systems. *Artificial Intelligence*, 20: 163-210, 1983.

CPHC. The Future Development of Computing in the United Kingdom. Workshop organised by the *Committee of Professors and Heads of Computing*, Manchester, January, 2000.

K. S. Decker and V. R. Lesser. Designing a Family of Coordination Mechanisms. *Readings in Agents* (eds. M. N. Huhns and M. P. Singh), Morgan Kaufmann. pp. 450-457, 1998.

J. E. Doran. Archaeological reasoning and machine reasoning. *Archaeologie et Calculateurs* (ed. J.C. Gardin) CNRS, Paris. pp 57-67. 1970.

J. E. Doran and F. R. Hodson. *Mathematics and Computers in Archaeology*. Edinburgh University Press. 1975.

J. E. Doran and M. Palmer. The EOS Project: integrating two models of Palaeolithic social change. *Artificial Societies: the Computer Simulation of Social Life* (eds. N Gilbert and R Conte), pp. 103-125. London: UCL Press. 1995.

H-P. Francfort. Modelling Interpretative Reasoning in Archaeology with the aid of Expert Systems: Consequences of a Critique of the foundations of Inferences. *Interpretation in the Humanities: Perspectives from Artificial Intelligence* (editors R. Ennals and J-C Gardin). LIR Report 71. The British Library. pp 101-129. 1990.

N. R. Jennings. Agent Based Computing: Perils and Promise. *Proc. 16th Int. Joint Conf. on Artificial Intelligence (IJCAI-99)*, Stockholm, Sweden. (Computers and Thought Award invited paper) 1429-1436. 1999.

N. R. Jennings and J. R. Campos. Towards a Social Level Characterisation of Socially Responsible Agents. *IEE Proceedings on Software Engineering*, 144 (1), 11-25. 1997.

A. C. Renfrew. Systems Collapse as Social Transformation: Catastrophe and Anastrophe in Early State Societies. *Transformations: Mathematical Approaches to Culture Change* (eds. A C Renfrew and K L Cooke) Academic Press. Pp. 481-506. 1979.

A. C. Renfrew and P. Bahn. *Archaeology: Theory, Methods and Practice*. Thames and Hudson. (2nd Edition). 1996.

M. Wooldridge, N. R. Jennings and D. Kinny. The Gaia Methodology for Agent-Oriented Analysis and Design. *Journal of Autonomous Agents and Multi-Agent Systems 3 (3)* (to appear) 2000.