Representing the Knowledge of the Expert Archaeologist

Richard Ennals
Derek Brough
Department of Computing
Imperial College of Science & Technology
180 Queen's Gate, London SW7 2BZ

To be presented at "Computer Applications in Archaeology 1982",
University of Birmingham, March 26-27.

The field of archaeology is now so extensive that no one individual can have complete knowledge of even one specialist area of work. Each researcher or enthusiast is dependent on the report and publications of field-workers, which will be subjected to the critical scrutiny of their academic peers prior to acceptance as reputable works of reference. The literature proliferates, often in a form not immediately usable by the subsequent enthusiast.

A number of problems lie behind this state of affairs. It is not generally agreed what an archaeologist is to do when he commits his expertise to paper. Is he to present a glorious catalogue of finds in the manner of Schliemann, a selective piece of autobiography following Howard Carter, another in one of the many series of introductions or field guides, a methodology of fieldwork in the style of Mortimer Wheeler, or simply a readable account, accessible to the layman, of the past history of a particular monument? Not only do we have to assess the nature of the expertise to be represented on paper or in some other medium, but the audience to which this expertise is to be communicated. There is a great difference between the form and purpose of a learned paper, and of a descriptive leaflet at an ancient monument.

It is arguable that conventional modes of publication are no longer adequate to satisfy the demand for the expertise of the archaeologist. The purpose of this paper is to suggest some alternative approaches, and concentrate particularly on the contribution to be made by logic programming and the development of expert systems.

What does someone look for when they consult an expert? Hawkins, in his "An Analysis of Expert Thinking" [Hawkins 1981] says "someone who can negotiate an agreed interpretation of a particular subject with the help of special knowledge and user opinions". He adds that "an expert appears very much as an analytical tool, helping the users make well-informed decisions without forcing them to accept any particular interpretation or procedure."

We are concerned in this paper with two different kinds of consultation of an expert: by the interested layman and by the specialist in a related subject domain. In the first case the questionner wants an answer in ordinarylanguage. In the second case the questionner wants to be able to communicate with other specialists and their areas of knowledge without being obliged to learn each other's specialised language and knowledge, through the medium of the expert.

Doran approached the problem of Knowledge Representation for Archaeological Inference [Doran 1977], "... in the light of the evolving artificial intelligence theories of specialist knowledge representation". We would agree with him that "archaeology has clear attractions as a problem-domain for artificial intelligence research", but would wish to advocate a different approach to the problem, arising from research in the logic programming community in recent years.

There has been a division in the Artificial Intelligence community over the appropriate form for representing knowledge. As Mylopoulos says in his overview of the area [Mylopoulos 1980], "the current design paradigm for 'intelligent' systems stresses the need for expert knowledge in the system along with associated knowledge-handling facilities" but "the basic problem of knowledge representation is the development of a sufficiently precise notation for representing knowledge". The fundamental division has been between the declarative and procedural approaches to knowledge representation. Doran's work falls into the procedural tradition, with knowledge conceived as a set of independent procedural or semi-procedural units. A simple approach would be to use production rules, of the form

"If A is encountered, then do x"

which is the approach of MYCIN [Shortcliffe 1978] and DENDRAL [Buchanan 1978]. Doran in his development of SOLCEM was more demanding, involving the use of a range of "recognition demons" that correspond to concepts rather than broad sources of knowledge. Mylopoulos observes that procedural schemes of knowledge representation allow the specification of direct interactions between facts thus eliminating the need for wasteful searching. However, he points out that it is difficult to understand and modify a procedural knowledge base.

By contrast, Mylopoulos emphasises the simplicity of notation of a declarative logical representation scheme, which leads to understandable descriptions and conceptual economy. On this approach a knowledge base is a collection of logical formulae which provides a partial description of a subject domain. There are considerable advantages in using predicate logic as the notation for knowledge representation. Not least of these is the fact that, following Kowalski's suggestion [Kowalski 1974] of the procedural interpretation of sentences of logic for problem solving, the language PROLOG has developed in which statements of declarative description can also be given a procedural semantics. More recently Hammond, Clark and McCabe have suggested that PROLOG is a suitable language for implementing expert systems, [Hammond 1980], [Clark, Hammond & McCabe 1981]. It is from this standpoint that we approach the problem of archaeological knowledge.

Beforeproceeding to the representation of archaeological knowledge we have to analyse its nature. Collingwood's remarks on the incomplete nature of the discipline of history also still holds true for archaeology: "The modern conception (of archaeology) as a study at once critical and constructive, whose field is the human past in its entirety, and whose method is the reconstruction of that past from documents written and unwritten, critically analysed and interpreted, was not established until the nineteenth century, and is even yet not fully worked out in all its implications", [Collingwood 1946]. Daniel reminds us [Daniel 1981] "that

our present state of knowledge is merely, like the present, a moment in time". Daniel maintains that the links with scientists and scientific techniques date from the beginning of this century. "By the outbreak of the war [1914] it was clear that archaeology as it developed was going to rely more and more on scientific techniques and was no longer the study of artifacts in the field and in museums in isolation. The archaeologist was becoming dependent on scientists who studied the flora, fauna and environment of his sites and who would date his material independent of archaeological methods." [P.147]. Mortimer Wheeler accepted this scientific component of archaeological research, but wanted to emphasize that "In a simple direct sense, archaeology is a science that must be lived, must be 'seasoned with humanity'". [Wheeler 1943].

Within the discipline of archaeology we can discern different kinds of knowledge. As Elias writes in his "Sociology of Knowledge" within such an activity "Not only knowledge of objects, but also knowledge of how to gain and how to advance knowledge of objects, how to catch them in one's net, and how to make nets, and how to make better nets, for catching them, develops over the generations." [Elias 1971, P.166]. We must not think that 'knowledge' is to be reduced to tabular data of some absolute kind. Cicourel observes [Cicourel 1964, P.108], "The actual analysis of tabulated data ... is dependent upon implicit theoretical and substantive knowledge obtained under considerably less rigorous conditions than that knowledge evident in the elegant tables presented".

It would be inappropriate to subdivide the task of archaeology into independent subtasks, for every act of archaeological interpretation is influenced by the results of every other act. Of course, however, archaeologists themselves simplify: the subject is broken down into specialist areas according to region or period or type of evidence. There will be factual knowledge that is particular to the specific domain, factual knowledge of the general subject area, knowledge of the procedures of the archaeologist, and general world knowledge deriving from the society and rationality within which he operates.

This amalgam of kinds of knowledge constitutes the expertise of the skilled practitioner, which is to be represented by our system. This assumes the existence of a unified activity or form of knowledge. The intention is that this should be modelled by a program, using the knowledge of an individual expert consultant, refining the model to generate appropriate correct output consistent with answers that would be given by the expert consultant himself.

1. Consultation by the interested layman

Our first example application of this approach to knowledge representation would be analogous to the introductory textbook, field guide or decriptive document available at museums and ancient monuments.

It would be argued that this was an appropriate area for an expert system on a number of grounds. There is a need to preserve the knowledge of expert curators whose services are being dispensed with as the Department of the Environment seeks economies and privatisation. The cost of microcomputers is falling while the level of public interest in archaeology is ever-increasing. Laymen need an introduction to diverse sources of

information, either in a formal classroom situation or in the form of references and information when interest has been aroused in a particular subject.

It is important that an expert system for the layman be easy to use. We use micro-PROLOG as our implementation language, a version of PROLOG written for microcomputers based on the Z80 microprocessor and the CP/M operating system [McCabe 1981], [Clark, Ennals & McCabe 1982]. We also make use of the sugared syntax of the front end "Simple" program in micro-PROLOG which renders the language into a more 'natural' form. This has also been used in developing materials for children [Ennals 1981] and for historical applications [Ennals 1981]. The mode of interaction is through querying and adding to a database. A list of names of relations used in the database is available, together with templates of the forms of the sentences in the relations and a 'help menu' of auxiliary commands.

The source of information for this example program was Wood's field guide "Archaeology in Britain" [Wood 1979]. The text can easily be represented in predicate logic in the micro-PROLOG form. Early on, abbreviations are provided for county names:

```
(Bedfordshire)      abb  Bd
(Berkshire)         abb  Bk
(Buckinghamshire)   abb  Bc
(Cheshire)          abb  Ch
```

Wood develops a glossary of useful technical terms:

```
(Barrow)          gloss   (mound covering a burial)
(Beaker people)   gloss   (first Bronze Age immigrants)
(Bell pit)        gloss   (pit for extraction of clay for iron-making)
```

He offers some generalisations regarding Patterns of Settlement:

```
x live ((Thames valley)) if x period (early Palaeolithic)
x live ((NE Wales) (Derbyshire) (Yorkshire) (E Coast)) if
                        x period (later Palaeolithic)
```

Most interesting perhaps is his section entitled "Identifying Earthworks" [P.77]. His account takes the form of a series of rules, each of the form:

```
If you see ...  it could be ...  For more detail, see Page ...
```

To enhance the compactness of the rules, further abbreviations are given, which we represent as

```
N     d    ((neolithic)    (4000 2400 BC))
B     d    ((Bronze Age)   (2400  700 BC))
E     d    ((early)        (2400 1500 BC))
M     d    ((middle)       (1500 1000 BC))
L     d    ((late)         (1000  700 BC))
I     d    ((Iron Age)     ( 700   50 AD))
R     d    ((Roman)        (  50  400 AD))
D     d    ((Dark Age))    ( 400  650 AD))
S     d    ((Anglo-Saxon)  ( 500 1100 AD))
Med   d    ((medieval)     (1100 1500 AD))
Mod   d    ((modern)       (1500 2000 AD))
```

Some of the rules can be represented as follows:

```
(low small round mound) could-be ((S Barrow) 154)
(group of low round or oval mounds) could-be ((S Cemetery) 155)
(conical mound often with flat top) could-be ((R Barrow) 153)
x possible-period y if x could-be ((z X) Y) and
                       z d y
x type y if x could-be ((z y) X)
x described-on-page y if x could-be (z y)
```

This information can be presented to the user in the form of a system that will identify a given earthwork, with the program questionning the user where further details are required to aid identification. A master program can call up and delete programs from disk as the interaction proceeds, making maximum use of core memory and of the facility for access to relations on disk.

2. Consultation by the specialist in a related subject domain

Hawkins [Hawkins 1981] is particularly concerned with the role of an expert system in aiding communication between specialists, each with their own technical language. He sees as the objective [P.6] "Each user experiences a simulation based on expert knowledge. A good simulation allows the expert to translate one user's hypothesis into results, expressed in another user's language. Such a translation exposes one specialist's opinion, via an expert's knowledge-based skill in simulation, to the scrutiny of another specialist". If this works, then "without being obliged to learn each other's specialised language and knoweldge, one user can communicate with another through the medium of the expert. The approach depends upon a sophisticated model of the understanding of the questioner: "The appropriate explanation would depend on the expert's assessment of what the questioner has failed to understand, i.e. the difference between the questioner's model and the expert's model." The expert remains a tool: "The expert's role is to assist one, or several, users to assemble mutually consistent models of the same geological object".

This kind of analysis could be applied to the construction of an account of what had happened to Tollund Man, found in a Danish peat bog. From the first discovery of the body by farm labourers, through the autopsy, the collection of archaeological evidence, the carbon-dating, the research into the customs of Germanic tribes, a number of specialists were involved in handling information with different technical terms, mediated by the expert archaeologists whose job it was to produce a coherent, consistent,

overall account. There were limits to what was possible. As Daniel wrote [Daniel 1981], "how much a reflection on archaeology is it that we could reconstruct his last meal but of course will never know his name, why he was hanged, or his last thoughts".

References

Buchanan, B.G. & Feigenbaum 1978 'Dendral and Metadendral: Their Applications Dimension', ARTIFICIAL INTELLIGENCE 11, 1978.

Cicourel, A.V. 1964 METHOD AND MEASUREMENT IN SOCIOLOGY, New York

Clark, K.L, Ennals, J.R & McCabe, F.G. 1982 A MICRO-PROLOG PRIMER, Logic Programming Associates, 36, Gorst Rd., London SW11.

Clark, K.L, Hammond, P & McCabe, F.G. 1981 'PROLOG: A Language for Implementing Expert Systems', MACHINE INTELLIGENCE 10, Eds. Hayes, Michie & Pao, Edinburgh University Press.

Collingwood, R.G 1946 THE IDEA OF HISTORY, Oxford Clarendon Press.

Daniel, G. 1981 A SHORT HISTORY OF ARCHAEOLOGY, Thames and Hudson.

Doran, J 1977 'Knowledge Representation for Archaeological Inference', MACHINE INTELLIGENCE 8, Eds. Elcock & Michie, John Wiley & Co.

Ennals, J.R 1981 'Logic as a Computer Language for Children: A One Year Course', DOC 81/6, Dept. of Computing, Imperial College.

Ennals, J.R 1981 'History & Computing: A Collection of Papers 1979-81', DOC 81/22, Dept. of Computing, Imperial College.

Elias, N 1971 'Sociology of Knowledge', SOCIOLOGY 5.

Hammond, P 1980 'Logic Programming for Expert Systems', M.Sc. Thesis, DOC 82/4 Dept. of Computing, Imperial College.

Hawkins, D 1981 'An Analysis of Expert Thinking', Schlumberger-Doll Research, Ridgefield, Connecticut. USA.

Kowalski, R.A 1974 'Predicate Logic as a Programming Language', IFIP 1974, Amsterdam.

McCabe, F.G 1981 MICRO-PROLOG PROGRAMMERS REFERENCE MANUAL, Logic Programming Associates, 36, Gorst Rd., SW11

Mylopoulos,J 1980 'An overview of Knowledge Representation', ACM PROC Workshop on Data Abstraction, Databases and Conceptual Modelling. eds Brodie M.L, Zillies,S.N.

Shortcliffe,E.H 1978 COMPUTER BASED MEDICAL CONSULTATIONS: MYCIN, North-Holland Elsevier

Wheeler,M 1943 MAIDEN CASTLE

Wood,E.S 1978 ARCHAEOLOGY IN BRITAIN, Collins Field Guide, Collins 1979 5th Edn.