Science and Archaeology no. 27 (1985) pp. 3-7

THE COMPUTERISED ARCHAEOLOGIST: THE DEVELOPMENT OF EXPERT SYSTEMS

Jeremy Huggett and Katharine Baker
Research Centre for Computer Archaeology,
North Staffordshire Polytechnic.

ABSTRACT

This paper contains a general review of existing expert system techniques, terminology, and a history of their development. After a detailed discussion of the methodology of knowledge elicitation and representation, the use of production rules in the knowledge database, search strategies, and uncertainty handling, the application of expert systems to archaeology is discussed. It is concluded that in their present form expert systems have a number of serious practical and theoretical problems. They will require a great deal of further work before being of use in archaeology.

INDEX

1. *Introduction*

The idea of a thinking machine has long been a staple of science fiction novels, films, and short stories. HAL in "2001: A Space Odyssey" is one example of a super-intelligent computer which (who?) has 'human feelings' while Deep Thought in the "Hitch Hiker's Guide" series is another. It is the ability of the bomb in the film "Dark Star" to reason which causes the problems - although its reasoning was obviously flawed!

On a less impressive scale, thinking machines are now becoming a reality, although much depends, of course, on your definition of thinking. Few would accept the proposition that computers can 'think' in the same way that humans can, but it has been demonstrated that in certain limited spheres of competence machines can already equal or surpass human performance. However, when the program of one of these 'intelligent' machines is examined, it is clear that the intelligence is only apparent and achieved by means of clever programming devices. What these machines actually do is to reproduce the results of what, in a human, would be considered to be intelligent thought. Confusion may arise because of the adoption of terms such as 'intelligence', 'reasoning' and 'understanding' which can give a misleading impression of these programs' capabilities.

Such programs have been used to make decisions and solve problems in a variety of areas and this paper is intended to discuss the development of these 'expert systems' and examine their potential and the problems connected with their use in archaeology.

2. *Some Definitions*

An expert system is fundamentally a 'problem-solver' program which operates most successfully within a restricted, pre-defined domain. There are two main features of an expert system. The first of these is the way in which knowledge is represented in the program, the second is the way in which it behaves towards the user.

The organisation of knowledge within an expert system is perhaps its main distinguishing feature when compared to a more conventional program. In a conventional procedural program, knowledge of a subject or domain is not held separately from the procedures:

> In such programs, knowledge about that application is scattered throughout the code, and changing a single fact may require a change to hundreds of lines of code in dozens of different modules (Sowa, 1984, p. 278).

In contrast expert systems employ a declarative approach, in which knowledge is held separately from the controlling program in a knowledge base. This knowledge base consists of a series of rules and facts which together provide a description of the subject domain. It may be modified as required, but any changes do not affect the program itself, simply the results that it gives. An expert system may be seen as having three main parts - first, the knowledge base; secondly, the inference engine or control system, which remains unchanged when the knowledge base is modified and which performs the reasoning of the system as a whole; and thirdly, a global database consisting of a store of information concerning the consultation at hand (Weiss and Kulikowski, 1984, p. 41).

The second distinguishing feature of an expert system is the way in which it behaves towards the user. The system is designed to use the domain knowledge represented within it to simulate the performance of a human expert in the same field. This simulation of expert performance requires that not only should the system be able to reason through a problem and provide a reliable answer, but that it should also be able to explain its steps in reasoning to the user, recognise if information is absent or untrustworthy and be able to request further information from the user. Thus the system may be interrogated about the questions that it has asked the user in the course of a consultation, or about the stage it has reached in its investigations, or about the conclusions it has offered, as well as providing accurate, reliable results - just as would be expected of a human expert.

Expert systems are seen as acting as consultants or advisors. Michie (1982, p. 197) divides this consultancy function into three "user-modes". The expert system may be approached by the user in the role of a client, seeking information and requiring diagnoses or prognoses. Alternatively the user may act as pupil, learning from the system through a series of questions and answers. Finally, the user may take the role of the tutor, refining, modifying or enlarging the system's knowledge. All three "user-modes" may be combined within a single system: indeed, a system would be much more powerful and be more likely to obtain user-acceptance if it did combine these features, rather than simply appearing to act as an omnipotent, omnicognisant oracle. One example of such a 'complete' system is that described by Davis (1982). He worked on a suite of three programs for medical diagnosis, each program effectively being an expert system but all sharing the same knowledge base (Davis, 1982, p. 58). MYCIN was the "performance program", supplying consultative advice to the physician; TEIRESIAS was the knowledge acquisition program, which allowed a human expert to educate the program; and finally, GUIDON functioned as tutor to a student. The system is therefore not a static program, but is open to verification, correction and expansion. Such a system and its three associated "user-modes" clearly has applications beyond that of medicine, and these characteristics should be seen as the sign of a 'mature' expert system. Not only are these three "user-modes" equally applicable to an archaeological expert system, they should also be seen as a requirement of any system that is to be actively used in the field.

Expert systems can also be divided into three main categories, defined by the problem areas to which they are applied: classification, design and decision support (Sowa, 1984, p. 280). Classification systems are by far the most common form of expert system. They handle a wide range of generally diagnostic problems, gathering data about a specific problem and suggesting a cause and/or a solution, and are the type used in medical problems and in mineral or oil prospection, for example. Such systems require large amounts of data, and may well be applicable to archaeological problems, many of which can be broken down into a diagnostic form (for example, the system for categorising beakers (Bishop and Thomas, 1984)).

Design-based expert systems are less likely to be of use to archaeologists at their present stage of development since they use exact reasoning on highly structured data. For instance, the expert system R1 (see, for example, Edelson 1982, pp. 50-56), configures components for the VAX line of computers produced by Digital Equipment Corporation. A design system will search for combinations that satisfy a given goal - given a set of components and their relationship to each other, it can design the layout of a computer chip, for example.

Decision support systems are seen by Sowa (1984, p. 280) as typically being intelligent front-ends to databases. They use the database and their knowledge base to explore alternatives, solve problems and make predictions, unlike a conventional database. These are often seen as being applicable to the business community, but they may also be of considerable use to archaeologists.

These divisions are not mutually exclusive: a single system may incorporate one or more features. However, for archaeological systems, one feature that is seen to be of crucial importance is the ability to handle uncertainty - not simply in the sense of incomplete data, but also in terms of imcomplete knowledge with which to reason about the data. This question of uncertainty will be discussed below.

The knowledge in an expert system is commonly represented as rules of the "If (A), then (B)" form, with facts simply being the trivial cases of rules, having nothing to prove.

3. *History*

Expert systems are an aspect of artificial intelligence which developed particularly from the late 1970s as part of mainstream artificial intelligence research.

The first attempts at thinking machines were based on the general exploration of human reasoning and tried to produce a few powerful solutions, applicable to any problem which could be reduced to mathematical arguments. It was hoped that the power of the computer coupled with representations of single laws of human reasoning would lead to an expert, and possibly even a superhuman performance from the machine. Various techniques were explored including theorem proving, searches, generate and test, and means-end analysis. Although limited success was achieved they failed to deal with real problems and performance was poor compared with that of a human expert.

The realisation that knowledge was as important to the problem-solving process as reasoning itself led to two major refinements. Domain-specific knowledge was incorporated into the programs and heuristic techniques were developed. Heuristic programming attempts to model the rules of thumb, informed guesses, and general domain knowledge which are used by human beings in problem solving. The introduction of domain-specific knowledge had three major benefits (Alty and Coombs, 1984, pp. 85-86): the real world problems which the general purpose problem solvers had been unable to handle could be tackled; the emphasis changed from the discovery of a few powerful rules to the representation of knowledge in such a way that it could be used effectively; and the nature of the processes which the programs were attempting to model were explored, that is, philosophical and psychological aspects were introduced. It was found that these systems were most effective in those areas where problem solving depended on personal knowledge and experience - hence the name 'expert systems' (Alty and Coombs, 1984, p. 86).

Among the earliest domain-specific systems were those applied to games such as chess (Michie, 1982) and chequers (draughts) (Samuel, 1963). These programs were also among the first to exploit the advantages of heuristic programming rather than relying on the simpler and cheaper 'brute force method' (Michie, 1982, p. 180). The latter is unselective and will examine all of the options available, continuing to do so even after reaching the point at which a human problem-solver would have abandoned a particular approach as being unlikely to lead to success. In theory this manner of problem solving will eventually produce the answer. In fact, however, because of the combinatorial explosion of possibilities in some problems the amount of processing time necessary becomes insurmountable. Boden (1977, p. 512) gives figures of 10^{40} possible board positions in the game of draughts, and 10^{120} board positions for chess, for example.

Early systems which used domain-specific knowledge included DENDRAL (Buchanan and Feigenbaum, 1978) which used a generate and test technique to outperform human experts in deducing the structure of organic molecules from mass spectrographs, and MACSYMA (Barr and Feigenbaum, 1982, vol. 2, pp. 143-

149) which excels in the field of differential and integral calculus.

MYCIN (Shortliffe, 1976) can be seen as the first of the fully-fledged expert systems. Not only does it incorporate domain-specific knowledge and exploit heuristic techniques, it also has an explanation facility, that is, it can explain its reasoning at the request of the user, and the knowledge base is separated from the program which manipulates it. The program identifies bacteria in blood and urine samples and recommends an appropriate therapy.

A number of expert systems are frequently mentioned in the literature and have now achieved the status of classics. In addition to those mentioned above this group includes PROSPECTOR, a system designed to aid in mineral exploration (Duda et al., 1979), and INTERNIST, now known as CADUCEUS, which carries out internal medical diagnoses (Weiss and Kulikowski, 1984, pp. 53-57).

Expert system shells may be developed by stripping the domain-specific knowledge away from the inference engine of an expert system. EMYCIN was constructed in this way using the MYCIN inference engine. In theory a new set of domain rules can be supplied and a new expert system formed. This has been done successfully using EMYCIN to produce PUFF, a system which performs analyses on lung function tests (Hayes-Roth, Waterman, and Lenat, 1973, p. 325). Other shells are now available such as APES which is an augmented PROLOG and includes a query facility (Hammond, 1982) and a new package from the Alvey directorate, which consists of several example shells for evaluation and assessment, is now on the market. Although shells have been claimed to be universally applicable, the type of inference engine inherited from the parent expert system limits the range of suitable applications. For example, EMYCIN is obviously most suited to those problems which can be tackled using diagnosis procedures.

Parallel with the growth of expert systems has been the development of languages particularly suited to their implementation. Much of the modern artificial intelligence work is carried out in high level languages such as LISP, LOGLISP, POP-11, POPLOG, and PROLOG. PROLOG, for example, has a backward chaining facility built into it. An expert system can be produced in any language, however, and many of the earliest artificial intelligence programs were written in languages such as FORTRAN and ALGOL (Boden, 1977, p. 35). Naylor (1983), for example, produced working expert systems using BASIC on a home micro computer. One problem with the specialised artificial intelligence languages mentioned above that they are mostly confined to mainframe installations and so may not be available to many archaeological users. A version of PROLOG, Micro-PROLOG, is available for some machines and was used, for example, by Brough and Parfitt (1984) to write a system which ages horse teeth.

4. *How they work*

There are three main areas of difficulty in the attempt to copy the reasoning of the human expert: knowledge elicitation, knowledge representation and search strategy. This is, how the domain knowledge is obtained, how it is structured and stored, and how it is used in the reasoning process. Hayes warns that we must clarify our expectations: do we wish to arrive at human-type conclusions (to emulate) or do we wish to reproduce the human-type processes (to model):

> it is perilously easy to conclude that, because one has a program which *works* (in some sense), its representation of knowledge must be *correct* (in some sense) ... if the program *works*, so the argument would go, then its representation must adequately capture the intended meanings, for that is what we *mean* by "adequate". (Hayes, 1979, pp. 244, 265)

Two main questions may be asked. The first is the extent to which these systems are epistemologically adequate, that is, whether the knowledge representation within such systems is sufficient. The second is whether these systems are heuristically adequate - is the reasoning mechanism employed reliable?

4.1 *Knowledge elicitation*

Before the stage of entering knowledge of a domain into a system is reached a problem may be encountered in the process of knowledge elicitation. One question is whether knowledge should be derived from "experts" or "practitioners" in a particular field, since there is a basic difference in the nature of the expertise which each of these groups has to offer (Hartley, 1981). Hartley suggests that if the knowledge that is to be represented is fragmentary and consists largely of facts which are additive in nature, then practitioners are more suitable, but if the knowledge is structural or systematic in nature, then experts are better suited (Hartley, 1981, p. 862). He argues that practitioners tend to act separately and therefore have different experience and knowledge, whilst the expert may have far greater experience than an individual practitioner but cannot match the aggregate experience of a group of them. On the other hand, experts have an ability to systematise a domain: "experts *are* experts because of their ability to see the structure of the domain and not merely its content" (Hartley, 1981, p. 862). A discussion of the character of archaeological knowledge with particular reference to the application of expert systems can be found in Reilly (forthcoming). Reilly considers that archaeology is primarily a craft-orientated discipline, and would therefore see practitioners as being best suited for most forms of archaeological knowledge elicitation.

Many philosophers have attempted to analyse the means by which a human expert approaches a problem, and there remains considerable disagreement. Some computer scientists working in the field often complain that they are faced with problems usually classified as being philosophical rather than purely practical (for example, Kayser, 1984, p. 168). Boden (1977), p. 125) claims that artificial intelligence studies have resulted in "a new standard of rigour, and a new appreciation of the importance of mental *process*", but this new appreciation is more a realisation of the complexities of human thought processes than any ability to actually reproduce more than their results. Few people are capable of the considerable degree of introspection required to accurately reproduce human 'knowledge'. As Sloman says:

> Many persons can recognise and use valid inferences even though they have learnt no logic and become incoherent when asked to explain why one thing follows from another ... (Sloman, 1971, p. 272)

There is a danger in attempting to explain the means by which a solution was reached since:

> we often invoke 20-20 hindsight and leave out the mistakes that we made along the way. Our explanation makes it appear that we followed a very direct and reasonable route from beginning to end (Stefik et al., 1982, p. 162).

Human reasoning employs elements of tacit inference:

> global knowledge of which one is not introspectively aware, but which usually determines the nature of the thought contents of which one is focally aware (Boden, 1977, p. 435).

All that knowledge representation in an expert system can achieve at present is a simulation of the results of a human expert's reasoning - in other words, knowledge engineers can construct a model of human reasoning which, given similar circumstances, will be able to reach conclusions which would correspond with those of a human expert. However, this model may achieve these results while bearing little or no resemblance to the human reasoning process, though it may attempt to emulate some of its characteristics, such as focused reasoning and 'best-guessing'. As Kayser (1984, p. 168) admits, these engineers' models are more concerned with robustness and efficiency than with completeness or consistency. Some systems, such as TEIRESIAS (see Davis, 1977, for example), assist in the transfer of expertise from human to machine, but the human expert has to conform to the system's model of reasoning and knowledge, rather than *vice versa*.

4.2 *Knowledge representation*

One of the changes observable in artificial intelligence work is the rise in importance of knowledge representation relative to that of search techniques (Feigenbaum, 1979, p. 7). The need is to represent knowledge in such a way that it can be used effectively. Mylopoulos (1980, pp. 8-10) gives four important ways knowledge representations in this area differ from those used in standard databases:

1) *Multiple use of facts* - facts stored in a knowledge base may well have more than one use. For example, a piece of information must be available for use by each inference mechanism used by the system.

2) *Incompleteness* - an artificial intelligence knowledge base is constantly developing. The world model does not, or at least should not, remain fixed as at the time of the original implementation. The knowledge base is not seen as complete: an answer "NO" from a system should be interpreted as "I DO NOT KNOW, I DO NOT STORE THAT INFORMATION".

3) *Self knowledge* - or "meta-knowledge", is knowledge about knowledge. The system may, for example, have guidelines as to which types of facts are useable in which types of situations. If this meta-knowledge can be represented in exactly the same way as the other facts in the knowledge base it is available for use in the same way. Meta-knowledge may be used in the knowledge acquisition process as it is, for example, in the TEIRESIAS system which is part of the MYCIN-GUIDON-TEIRESIAS suite (Davis, 1982). Another aspect of self knowledge is a system's ability to answer questions about its actions and lines of reasoning put to it by a user, this ability being seen as an important feature of an expert system by many commentators.

Mylopoulos (1980) has divided schemes of knowledge representation as follows, although he notes that few implemented schemes are actually this simple. The main distinction made is that between declarative schemes and procedural schemes. PROLOG attempts to combine the advantages of both types. Declarative schemes can be divided into logical schemes, and network or semantic schemes. Procedural schemes can be divided in terms of the ways in which they deal with control structures and also with the triggering of the procedures themselves. The use of "frames" for representing knowledge employs different ideas and is usually treated separately.

4.3 *Declarative schemes*

Logical representation schemes represent their facts as a collection of logical formulae. The formulae may be composed according to the rules of any suitable logic. Their advantages include a sound theoretical grounding in the particular logic used, a relatively simple and therefore easily understood notation, and their suitability for use in many types of inference mechanisms. Among their disadvantages are the lack of established rules for organising the formulae within a large knowledge base and the difficulty of representing heuristic or procedural knowledge.

In a network or semantic scheme the domain is represented in terms of objects and the associations between objects. There is a strong link between this form of representation and the problems of search strategies as the associations between the objects can actually be used to define access paths through the knowledge base. A second advantage is that the knowledge base can be more easily organised than the logical representation while the notation can be reproduced in the form of a graph and so is easily understandable. However, there is a considerable variety of ways in which networks have been used and hence a lack of an agreed terminology.

4.4 *Procedural schemes*

Many procedural schemes are implemented in LISP or in LISP-type languages. The knowledge base is simply a collection of procedures. Production systems are a subset of the procedural representation group and are probably the most familiar form of knowledge representation. Mylopoulos (1980, p.8) identifies one major advantage and one major disadvantage that procedural representations have when compared with declarative schemes: while they can cut out unnecessary searching by specifying the interactions between facts, the actual knowledge base is hard to understand and difficult to modify. Schemes based on this representation do, however attempt to rectify the disadvantages.

4.5 *Frames*

The basic unit of this type of representation, the frame, is also known as a "script", or a "prototype". In this type of representation the knowledge base consists of a series of frames, each frame corresponding to a different stereotyped situation and including information on, for example, key objects, relationships, and default values. The scheme was originated by Minsky (Minsky, 1975). The first proposals were simple and used ideas from most of the other representation schemes. Later versions have developed different aspects: FRL uses a hierarchical organisation for the knowledge base, KRL possesses a degree of self-knowledge and allows for a focusing on the problem in hand. Full references and further examples can be found in Mylopoulos (1980, p. 8).

4.6 *Production systems*

The problem of actually extracting the knowledge from a human expert is exacerbated by the way in which most expert systems represent that knowledge. Not only is it unlikely that human experts reason along production system lines, but many specialists consider that the type of production system used in most expert systems to date is unsuitable, or at least inadequate, as a means of knowledge representation. Davis has argued for example that rules are only appropriate to a particular form of knowledge, and that it is possible that:

> the knowledge in some domains is inherently unsuited to a rule-like representation, since rules become increasingly awkward as the number of premise clauses increases (Davis, 1982, p. 76).

In other words, production systems are inadequate representations of complex systems which may work well in limited areas but which will soon break down. Davis (op. cit.) suggests that knowledge at a certain level of formalisation may be appropriate for the application of production rules: before this stage, the knowledge is too unstructured while after it, straightforward algorithms may be more applicable. Just where archaeological knowledge fits into this scheme is unclear, although an intuitive judgment would suggest that it is for the most part highly unstructured. Some have obviously considered archaeological knowledge in certain limited areas to have reached a level at which some elements can be represented as rules (Brough and Parfitt, 1984; Ennals and Brough, 1982, for example). The limitations of production systems and similar existing means of representation are recognised by some (for example, Sloman, 1979, p. 240), but more powerful descriptive formalisms have yet to be developed. Until then, a common attitude is often that if it appears to work, the representation must be adequate. Hayes (1979, pp. 244, 265) warns against this behaviouralist criterion on the grounds that, although the system may appear to work, the compromises and simplification required to enable the program to operate results in a sparser, more limited and less thorough representation which will clearly have repercussions on the performance and reliability of the system.

The amount, or more specifically, the *level* of the knowledge represented within a system will fundamentally affect the way in which it performs. Two general levels can be isolated - 'shallow' and 'deep'. Shallow systems contain only enough knowledge to be able to perform a particular task adequately (whatever that criterion of 'adequacy' might be). In order to do this, such systems are limited to restricted, well-structured domains which are treated as being largely self-contained, and they only operate within these bounds. Most expert systems which have been produced fall into this category. A major problem that arises with such programs is known as the "plateau and cliff effect", and results in non-uniform performance.

> the program is outstanding on the core set of anticipated applications, but degrades rather ungracefully for problems just outside its domain of coverage. On very difficult cases, which are not typical of the ones used in formal evaluations, the programs may even be misled in cases that fall within their central domain by complex interactions or multiple disorders that they are unable to untangle successfully (Szolovits, 1982b, p. 15).

Given the emphasis placed by artificial intelligence experts on restricted domains as a requirement for a successful expert system (for example, Stefik et al., 1982, p. 142; Barr and Feigenbaum, 1982, p. 183) this situation is unlikely to improve. Szolovits (1982b, p. 15) argues that a system cannot be expected to be flawless when the domain of expertise is subject to uncertainties of knowledge and lack of data which might equally well trip up a human expert. Needless to say, an extensive period of evaluation and modification ought to extend the plateau of a system's performance and this may be further extended by the addition of a wider (but consequently larger, more detailed and complex) knowledge base in an attempt to provide depth to the program's knowledge.

A 'deeper' system not only requires a larger knowledge base, but that knowledge should also consist of both static knowledge - the rules and facts - and knowledge about the concepts that resulted in the choice of those rules and facts and which determined the methods of inference employed. To some extent, therefore, 'deep' systems should be able to reason *about* their knowledge, not just with it. Doran, for example, saw the inclusion of concepts as well as knowledge about the domain as being important for archaeological inference (Doran, 1977, p. 453). Arising out of this, another area in which a deep expert system could improve on the performance of the shallow systems is in the handling of conflict. As Szolovits says:

> conflict, just as agreement, is reduced to a manipulation of strength of belief. Yet, by contrast, we believe that human experts make a much more powerful use of occasions where they detect conflict. They are not satisfied by a simple revision of their degree of belief in the hypotheses which they have previously held; they seek a deeper, more detailed understanding of the causes of the conflict they have detected ... Conflicts provide the occasion for contemplating a needed re-interpretation of previously-accepted data ... and the reformulation of hypotheses ... (Szolovits, 1982b, pp. 16-17).

This kind of ability would considerably increase the power of expert systems, and, if developed, could result in the ability to test and develop new hypotheses. Such a system ought to be able to suggest relevant avenues of approach if unable to supply a solution to a problem, or point out parallels and analogies between the existing situation and others that it has dealt with in the past, for example. Such expertise in a machine would represent a far greater challenge to a human expert than do the existing shallow systems.

A further step that has yet to be taken but which a number of leading artificial intelligence practitioners have argued is necessary for the construction of properly competent and reliable programs, is the incorporation of informal expertise, or common-sense, in expert systems. The indefinable ability of humans to assess whether an approach is 'sensible' or not is based on a considerable body of experience which is largely tacit in nature and may often appear to be guesswork. Such a facility in an expert system would enable greater reliability of reasoning within the model of formal expertise - the rules and facts in the knowledge base. Whether such a development is altogether desirable is another matter - it might be seen as a vast leap towards the ultimate takeover by ultra-intelligent machines spoken of by Michie (1982, pp. 225-226).

4.7 *Search strategies*

Search strategies control the reasoning of an expert system and are handled by the inference engine. The most obvious division is that between general control methods and specific search methods. General control methods can be further divided into forward and backward chaining although this should not suggest that the techniques are mutually exclusive. MYCIN, for example, uses a backward chaining inference engine but also uses antecedent rules which proceed by forward chaining.

4.7.1 *Forward chaining*

This strategy starts from known facts and works forwards towards a conclusion. A premise is matched with the antecedents of a rule the consequents of which are then fired. Rules employed in this way are sometimes known as "demons" - for example, Doran (1977) proposed the use of "recognition demons" in his system for cemetery analysis SOLCEM-D. A problem with this method is the process may become hard to control as the number of routes to be explored increases rapidly leading to a combinatorial explosion: matches may be made with a large number of antecedents, which although irrelevant to the actual problem, will be followed up until a match fails.

4.7.2 *Backward chaining*

This strategy moves from the consequents of the rules to their antecedents. The premise is matched to the consequent of the rule, the antecedent of which is then treated as a new premise. A backward chaining inference engine may be seen as "guessing" at a solution and then attempting to prove it. If the attempt fails the system will then "guess" at another solution. The technique provides a more focused pattern of reasoning than forward chaining.

Searching may also be controlled more immediately by the use of heuristic weightings. These weightings are applied to rules to influence the system's choice between potential solutions. These weightings have also been used to indicate the level of confidence which the user can place in any solution, and also to indicate the amount of evidence for (positive) and against (negative) any statement. They are, therefore, connected with the major problem of handling uncertain evidence.

4.8 *Handling uncertainty*

One of the main ways in which human reasoning differs from that carried out by machines, at least so far, lies in the degree to which satisfactory and relevant conclusions can be drawn from uncertain, incomplete or contradictory evidence, where the situation is extremely complex or in situations where there are added complications such as lack of time. These abilities are used not only in day-to-day common sense decisions but also in those areas seen as the territory of the expert. These problems may well arise because a human expert finds precision in the situation unnecessary, or because there are no suitable units of measurement for the variable (Zadeh, 1979, p. 158). It may even be that it is difficult to define meaningful variables. Most of the successful expert systems which have been produced so far have been based upon well structured domains in areas with a tradition of "scientific" thought: for example, medical diagnosis, geology, mass spectroscopy, differential and integral calculus. It is particularly in those areas of knowledge which do not have this orderly infrastructure that the problems arise: over-simplification; the forcing of knowledge into forms of representation which constrain or distort it unrealistically; failure to exploit the human expert's vast stores of heuristic knowledge; and the difficulty of extracting knowledge from "tame" experts when they may be unable to express adequately what it is that they know or how it is that they reach their conclusions. Archaeology is a prime example of a subject in this position. The problems of elicitation are recognised in the "scientific" domain areas and are likely to be even greater in the humanities subjects with their traditional distrust of the dehumanising effects of anything seen as "scientific". The resolution of these problems is an important research topic.

Several different methods for coping with uncertainty have been put forward, although no one method has achieved pre-eminence, or even a general consensus as to its validity. Some, such as that employed by MYCIN, have been proved to work in their real environments although open to theoretical and practical criticism (Cendrowska and Bramer, 1984). Bearing in mind the point that different domain applications are likely to need different representations and search strategies if they are not to be unrealistically altered, this is not surprising.

If it were possible to store all possible patterns of evidence from within a domain and also to store all the conclusions which could be generated from that evidence, a system could emulate a human expert simply by means of consulting a look-up table (Weiss and Kulikowski, 1983, pp. 26-27). Not only is this undesirable in terms of computer time and facilities, it is common for the same evidence patterns to lead to different conclusions. The nature of some domains does not permit such an analysis: some domains cannot be reduced to mathematical formulae. Even if some means of judging between options were found the combinatorial explosion would be unmanageable. The efffort to find a satisfactory combination of heuristics and a means of deciding between options is a main area of research.

An obvious method of making such a decision would be to assign weightings to each fact or rule in the knowledge base. In analysing the human expert's methods of decision making it becomes obvious that they tend to be informal and intuitive. The lack of statistically adequate samples for many domains means that the human must be relied upon to provide these weightings. Standard statistics and probability theory may be able to deal with many of the resultant problems if the number of hypotheses and the size of the knowledge base is not too great (Duda, Hart and Nilsson, 1976, p. 1). Systems may use pure or modified Bayesian theory as the basis for their weightings, or they may use an individual ad hoc method. MYCIN uses its own system of two values for each proposition, one for evidence supporting the proposition, and one for evidence denying it. These figures are not probabilities as they are independent of each other (Quinlan, 1982, p. 9).

The work of Zadeh (for example, 1979) is strikingly different from the other systems with its use of linguistic quantification. He bases his ideas on the concept of a possibility distribution rather than a probability distribution, that is on fuzzy logic, and sees them as being particularly relevant to those areas where standard logic is unsuitable, where judgments are qualitative rather than quantitative. His possibility distribution is based on linguistic divisions (for example, true, false, not true, not very false, more or less false) which are expressed in the natural language meaning representation language PRUF (Possiblistic Relational Universal Fuzzy) (Zadeh, 1979, p. 149).

The nature of archaeological knowledge suggests that the handling of inexact information is likely to be important in an archaeological expert system. A similar need to handle uncertainty has arisen in other fields and a number of decision-making techniques have been developed for use in expert systems. Szolovits and Pauker (1978) identify two classes of reasoning or decision-making - categorical and probabilistic. Categorical decision-making uses a set of routines or rules which apply most, if not all, of the time. They need not be absolutely deterministic (as for example, the reasoning in the expert system R1) but are appropriate to most straightforward situations.

> A categorical decision typically depends on relatively few facts, its appropriateness is easy to judge, and its result is unambiguous. A categorical decision is simple to make and the rule which forms its basis is usually simple to describe (although its validity may be complicated to justify) (Szolovits and Pauker, 1978, p. 117).

This type of reasoning may employ rules-of-thumb which are true most of the time. Thus, an archaeological generalisation such as 'if two graves intersect, they are likely to be well-separated in time' is most likely to hold true, although there will doubtless be some occasions when this is not the case. However, such reasoning on its own is inadequate for anything other than a highly structured domain, and while archaeologists may employ categorical reasoning at times, it is insufficient to cope with all situations. Categorical reasoning cannot be used in situations where there are a number of factors, some of which may be interdependent, many of which may be uncertain, and all of which may have some bearing on a particular decision. Probabilistic reasoning attempts to emulate the human ability to assess situations and weigh evidence for and against a hypothesis or action. This approach to the confirmation or otherwise of hypotheses has also been proposed for use in archaeological theory-building (see, for example, Salmon, 1982). Most expert systems employ variations on the theme of Bayesian decision theory and consequently all are subject to its limitations to some degree. It should be stressed that these limitations are inherited from the use of Bayesian approach, and some systems, such as MYCIN, have attempted to circumvent some of these problems.

The Bayesian approach is limited by its appetite for data. A direct implementation of Bayesian decision theory requires a large database - for example, a relatively small problem consisting of 10 possible hypotheses, 5 possible tests, each with 2 possible results, would require 63,300 conditional probabilities (Szolovits and Pauker, 1978, p. 120). As a result, independence assumptions are generally made in order to reduce the amount of data required and to make calculation easier. The assumptions made are that the order in which the possible tests are performed is unimportant and that each test is independent of every other test - the result of one test does not rely upon the results obtained from any other test. This considerably reduces the data required: the example given above is reduced to a total of approximately 100 for instance. In addition, Bayesian theory assumes that there is a fixed finite set of exhaustive and mutually exclusive hypotheses within a problem domain. However, in archaeological terms at least, there can never be assumed to be an exhaustive set of hypotheses, let alone a set in which they are all mutually exclusive. Nor can the independence assumption be safely made; in fact, all the basic assumptions of Bayesian decision theory are too simplistic in nature and can be shown to be quite false in many situations. The result of this is that a system using such a technique may give conclusions or predictions that are quite simply wrong. Similar problems remain in the MYCIN expert system in spite of its different method of handling uncertainty (see, for example, Cendrowska and Bramer, 1984, p. 489).

Probability-related methods of dealing with uncertainty all reduce likelihood to a numerical value, and this gives rise to another problem - the reliability of the values that are assigned to hypotheses, be they called confidence factors, degrees of belief or probabilities. It has often been stated that experts do not easily carry probabilities in mind, and while they may *feel* that an answer is more or less likely, they are unable to assign a reliable numerical value to that feeling. For example, Szolovits and Pauker discovered that their medical colleagues:

> are often extremely reluctant to engage in any numerical computation involving the likelihood of a diagnosis or prognosis for a treatment. Even when official blessing is bestowed upon Bayesian techniques, we have seen both experienced and novice physicians acknowledge and then ignore them. Doctors certainly have a strong impression of their confidence in a diagnosis or treatment, but that impression must arise more from recognising a typical situation or comparing the present case to their past experiences, rather than from any formal computation of likelihoods (Szolovits and Pauker, 1978, p. 142).

McCarthy and Hayes go further in their criticism of these techniques:

> The information necessary to assign numerical probabilities is not ordinarily available. Therefore, a formalism that required numerical probabilities would be epistemologically inadequate (McCarthy and Hayes, 1969, p. 490).

One of the problems in assigning numerical judgments to hypotheses is that the assignment of one probability implicitly assigns a probability to the opposite condition. Thus a rule which has a weighting of 0.7 implicitly states that there is a 30% chance that the opposite may be true. While a specialist may consider 0.7 to be a reasonable assessment of his confidence, he will often disagree with the hidden probability of the opposite conclusion (Weiss and Kulikowski, 1984, p. 28).

There is also the question of consistency to be considered. The assignment of a numerical measure of confidence will differ not only from person to person: a single person may lack consistency through time and even during the course of one consultation. These values can only be approximations: the means of achieving them is basically unsound, and therefore they may actually add to the uncertainty that they were in fact intended to alleviate. Even techniques of fuzzy logic are subject to the same criticisms - the prior assignment of values may be fundamentally unreliable, and in the case of fuzzy logic the problem is compounded since the assumptions that were made in their formulation tend to be concealed more effectively. In other words, numerical values are unreliable measures of confidence, and however reliable the statistical techniques that manipulate them may be (and these have problems of their own), the actual confidence that can be placed in these values is small. It seems unlikely that archaeologists will be more successful in assigning reliable levels of confidence to their hypotheses than any other group, and so these techniques for handling uncertainty are clearly heuristically inadequate for archaeological purposes at least. Having said that, it has to be admitted that the often *ad hoc* schemes that have been used in systems such as MYCIN, DENDRAL and PROSPECTOR have performed with some success, and in fact Bayesian decision theory may prove to be adequate in very restricted problem domains (Szolovits and Pauker, 1978, p. 122).

5. *The Representation of Archaeological Knowledge*

Expert systems and the models of domain knowledge that they contain are incompatible with many elements of archaeological knowledge. Archaeological knowledge, indeed, all knowledge, is a continuum, always changing through time at varying rates. At any stage, knowledge may be abstracted to form a discrete model, but that model can only be an approximation to the continuum as a whole:

> For any specific purpose, a discrete model can form a workable approximation to a continuum, but it is always an approximation that must leave out features that may be essential for other purposes. Since the world is a continuum and concepts are discrete, a network of concepts can never be a perfect model of the world ... A closed, rigid system maintains a sense of security by giving instant answers to all perplexities. But it is a false security that is threatened by any incompatible viewpoint. (Sowa, 1984, p. 345).

Archaeologists already abstract and simplify in order to construct discrete models, but these models can be freely discussed, modified, accepted and rejected, while arguably those models within a machine will be less accessible for these processes. More importantly, the requirements necessary for the construction of an expert system model world may have considerable repercussions on the actual form of the knowledge contained within it. For example, knowledge representation within an expert system demands a high degree of formalisation, and while artificial intelligence specialists disagree about the way in which the knowledge should be represented, archaeologists argue about whether formalisation is desirable or not. According to Weizenbaum, abstraction involves:

> leaving out of account all those empirical data which do not fit the particular conceptual framework within which science at the moment happens to be working. (Weizenbaum, 1976, p. 127).

Expert systems present the possibility of actually fossilising the "particular conceptual framework" that was current at the time of abstraction. Whilst the knowledge base may be changed, the actual theory that governed the original selection of the rules and facts contained within it, and which resulted in the initial abstraction in the first place, remains hidden from the user. Systems that currently operate are ultimately unable to justify their model of a domain in terms of the assumptions that were implicit at the time of abstraction (for example see Clancy, 1983). This lack of depth in a system could be of critical importance, particularly if that system was being used by a pupil or non-specialist, in which case incorrect or inapplicable advice might go unnoticed.

The actual nature of archaeological knowledge also differs from the type of knowledge that is represented in expert systems. Expert systems deal in relationships between the rules and facts in the knowledge base and those relationships are basically very simple. The associative or pattern-matching inference engine found in most expert systems operates blindly, without comprehension and with only a relatively crude controlling mechanism in the form of heuristic weightings. The inference engine:

> takes into account simply the form of the particular problem, the theorem to be proved or the question to be answered, and whatever is relevant from the knowledge base, again by virtue of its form and not its content. The machine does not know or care what you are talking about. It covers only whether what you say follows from what you have agreed to assume (Ennals and Briggs, 1984, p. 135).

Expert systems have therefore been applied on the whole to areas where the knowledge is well-structured and capable of being reduced to a logical system of relationships. 'Reduced' is perhaps the term to note here - the process of abstraction is fundamentally one of reduction. Some areas of archaeological knowledge may well be capable of sustaining this level of formalisation, but many areas are clearly not at present and some may never be. Regardless of whether formalisation is desirable or not, archaeological knowledge consists of far more than straightforward logical relationships. Expert systems basically handle phenomenological knowledge; that is, they are:

> concerned with the relations among phenomena more than with an understanding of the mechanisms which are suggested by the observations (Szolovits, 1982, p. 16).

Much of what constitutes archaeological knowledge surely falls into the category of "understanding mechanisms which are suggested by observations". Logico-deductive methods of reasoning also have their limitations (see Salmon, 1982, pp. 38-9, for example). Logic can confirm or deny a hypothesis or prediction, but it can do little else. "Using solely deductive logic, one can only formulate the alternative hypothesis that is the denial of the original hypothesis" (Salmon, 1982, p. 55).

Clearly, therefore, existing types of systems have inadequate forms of knowledge representation, and while new techniques are being developed, these remain strictly experimental and still require a large amount of structure and formalisation in the application domain. These problems are not unique to archaeological applications, but the effects which the use of such systems may have on archaeological knowledge are considerable.

Given the problems described above, it would seem to be clear that knowledge representations are basically epistemologically inadequate. That alone should not mean that such systems ought not to be implemented, but it does mean that any archaeological system should be considered as being primarily experimental in nature. Artificial intelligence experts are well aware of the problems of knowledge representation (see, for instance, Newell, 1982) but as archaeologists, we should be wary of systems developed by non-archaeologists who simply see archaeology as a convenient test bed and archaeologists as receptive and naive users. "At the moment, it would be foolhardy to place all our archaeological eggs in the knowledge engineers' basket!" (Reilly, forthcoming).

6. *The Present Use of Expert Systems in Archaeology*

The application of expert systems in archaeology is a relatively new phenomenon. A number of uses have been proposed, but few have been fully implemented and none used in the field.

Doran (1977) recognised the potential advantage of knowledge-based systems over procedural schemes for archaeological representations. His procedural SOLCEM program for analysing cemetery test data was fundamentally limited by its lack of actual knowledge of cemeteries (Doran, 1977, p. 444), a problem which was approached in his proposed SOLCEM-D system by employing a series of knowledge-bearing units, called 'recognition demons' which acted independently of each other on an evolving interpretation structure (Doran, 1977, pp. 446-447).

Ennals and Brough (1982) suggested a somewhat different approach as a result of their criticism of Doran's outlined system. They saw Doran's SOLCEM-D as remaining within the procedural framework, with knowledge represented as sets of procedural or semi-procedural units (Ennals and Brough, 1982, p. 57). Instead, they proposed the use of declarative logical representation schema, of the type offered by the language PROLOG, where knowledge was broken down into declarative production rules, as found in the MYCIN and DENDRAL systems for instance. Doran had dismissed production systems as being too simple for archaeological representations (Doran, 1977, p. 445), but Ennals and Brough argued that archaeologists already themselves simplify, and break the subject down into specialist areas which are then further subdivided (Ennals and Brough, 1982, p. 58). Ennals and Brough proposed a system, implemented in Micro-PROLOG, which acted as a field guide to archaeological monuments, particularly the identification of earthworks (Ennals and Brough, 1982, pp. 59-60).

Bishop and Thomas (1984) described the implementation of an expert system on a BBC micro which attempted to encapsulate the knowledge represented by Clarke's classification of Beaker pottery. The system was effectively a computerised typology, based on Clarke's categories, and isolated a series of characteristics which allowed it to state a probability that a given beaker belonged to a particular category, creating additional categories if required. A similar system was described by Brough and Parfitt (1984), designed to age horse remains from their teeth, and Bourelly and Chouraqui (1984) produced a classification system for Mediterranean wine amphorae.

A number of observations may be made concerning the systems proposed by Ennals and Brough (1982), Bishop and Thomas (1984) and Brough and Parfitt (1984). These systems are all effectively classification systems, collecting the characteristics of an object, and producing a diagnosis or identification. From the published accounts, none of the systems employed any form of probabilistic reasoning - straightforward logical reasoning appears to have been the technique used. All the systems are based on well-structured data - a list of rules associating earthwork shape with date (Ennals and Brough, 1982, p. 60), the categories and criteria laid down by Clarke for classifying Beaker pottery (Bishop and Thomas, 1984), and a knowledge base derived from a well-established set of zoological criteria (Brough and Parfitt, 1984). None of these systems attempt to encapsulate 'concepts' as well as 'knowledge' as Doran required for SOLCEM-D, but they simply operate on the level of practical identification or interpretation. This is not to imply any major criticism of these systems: classification systems have been proved to perform well on a variety of knowledge bases, but the theory behind these systems remains opaque to the user - Doran's vision of knowledge-based systems having an impact on the conceptual and theoretical levels of archaeology (Doran, 1977, p. 453) has not yet been realised. As a result of this, and also as a consequence of the general overselling of expert systems, dissatisfaction with these forms of system has been increasingly expressed (see, for example, Wilcock (forthcoming) for a critique of expert systems and their archaeological applications).

A 'second generation' of archaeological expert systems is being developed. One project (Baker) is designed to establish a classification-type system using taxonomic data, developing it further into a 'deeper' system which incorporates knowledge of the concepts and techniques of environmental archaeology, in order to produce reasoned strategies for sampling and to make inferences concerning, for example, diet. Another project (Huggett) is not concerned with the construction of a classification-type expert system, but is examining the development of a system which effectively acts as an intelligent front-end to an already existing database, and uses the data to test hypotheses, explore alternative explanations and display expertise at the more theoretical levels of archaeology rather than simply performing practical identifications of artefacts.

7. *Conclusions*

Artificial intelligence programs in general, and expert systems in particular, pose a number of problems for any domain or discipline which may wish to use them. Indeed, Boden (1977) likens the artificial intelligence community to the Sorcerer's Apprentice:

> The apprentice learnt just enough magic from his master to save himself the trouble of performing an onerous task, but not quite enough to stop the spellbound buckets and brooms from flooding the castle (Boden, 1977, p. 463).

In their current state of development, expert systems certainly do not have all the answers to all the problems. There is a danger, however, that once it is in operation, users of such a system may come to see it not as the apprentice but as the wizard himself. Indeed, expert systems seem often to create more problems than they solve, although this may not be realised by designers who may not fully appreciate the problems of the application domain. Some of the features or requirements of expert systems may in fact be wholly incompatible with the structure and organisation of a particular domain's knowledge, although the very need for rigidity of structure and rigour of thought which these systems impose can sometimes be seen as positively beneficial to the host discipline.

At first sight, expert systems offer a great deal to an overworked and under-funded archaeological community. However, caution must be exercised, since at present, as has been discussed, there are considerable problems in knowledge representation and reasoning which will affect their performance and hence their acceptability. One danger is that archaeology may become a test bed for artificial intelligence research which could result in the development of systems that are archaeologically inadequate. The adequacy or otherwise of archaeological expert systems can be judged only by archaeologists. Expert systems should not be given a permanence or status that they do not deserve. By its very nature, an expert system which is perceived as functioning adequately - that is, producing the sort of answers that are expected - will cease to be archaeologically adequate once the conceptual framework under which it was developed has changed. In essence, expert systems are less flexible than the humans they attempt to emulate.

REFERENCES

Alty, J.L. and Coombs, M.J., 1984
Expert Systems: Concepts and Examples (NCC Publications).

Barr, A. and Feigenbaum, E.A.,1982
The Handbook of Artificial Intelligence (Pitman).

Bishop, M.C. and Thomas, J., 1984
"BEAKER" - An Expert System for the BBC Micro *Computer Applications in Archaeology* pp. 56-62.

Boden, M., 1977
Artificial Intelligence and Natural Man (Harvester Press).

Boden, M. 1984
Methodological Links between AI and Other Disciplines in Torrance, S. (ed.) 1984, pp. 125-32.

Bourelly, E. and Chouraqui, E., 1984
Systeme Experte et Simulation d'un Raisonnement en Archeologie *Informatique et Gestion* 151, pp. 46-51.

Brodie, M.L. and Zillies, S.N. (eds), 1980
Proceedings of Workshop on Data Abstraction, Databases and Conceptual Modelling (ACM).

Brough, D.R. and Parfitt, N., 1984
An Expert System for the Ageing of a Domestic Animal *Computer Applications in Archaeology* pp. 49-55.

Buchanan, B.G. and Feigenbaum, E.A., 1978
DENDRAL and MetaDENDRAL; Their Applications Dimension *Artificial Intelligence* 11(1), pp. 5-24.

Cendrowska, J. and Bramer, M., 1984
Inside an Expert System: A Rational Reconstruction of the MYCIN Consultation System in O'Shea, T. and Eisenstadt, M. (eds), 1984, pp. 453-97.

Clancy, W.J., 1983
The Epistemology of a Rule-Based Expert System - A Framework for Explanation' *Artificial Intelligence* 20(3), pp. 215-51.

Davis, R., 1977
Interactive Transfer of Expertise: Acquisition of New Inference Rules *International Joint Conference on Artificial Intelligence* pp. 321-28.

Davis, R., 1982
Consultation, Knowledge Acquisition, and Instruction: A Case Study in Szolovits, P. (ed.), 1982a, pp. 57-78.

Doran, J., 1977
Knowledge Representation for Archaeological Inference, *Machine Intelligence* 8, pp. 433-54.

Duda, R., Gaschnig, J. and Hart, P., 1979
Model Design in the Prospector Consultant System for Mineral Exploration in Michie, D. (ed.), 1979, pp. 153-67.

Duda, R., Hart, P. and Nilsson, N., 1976
Subjective Bayesian Methods for Rule-Based Inference Systems (Technical Note 124, A.I. Centre, SRI International).

Edelson, E., 1982
Expert Systems - Computers that think like people *Popular Science* (Sept. 1982), pp. 48-52.

Ennals, R. and Brough, D.R., 1982
Representing the Knowledge of the Expert Archaeologist *Computer Applications in Archaeology* pp. 56-62.

Ennals, R. and Briggs, J., 1984
Logic and Programming in Torrance, S. (ed.), 1984, pp. 133-44.

Feigenbaum, E.A., 1979
Themes and Case Studies of Knowledge Engineering' in Michie, D. (ed.), 1979, pp. 3-25.

Feigenbaum, E.A. and Feldman, J., 1963
Computers and Thought (McGraw-Hill).

Hartley, R.T., 1981
How Expert Should an Expert System Be? *Proceedings of the Seventh International Joint Conference on Artificial Intelligence* vol. 2, pp. 862-67.

Hayes, P., 1979
The Naive Physics Manifesto in Michie, D. (ed.), 1979, pp. 242-70.

Kayser, D., 1984
A Computer Scientist's View of Meaning in Torrance, S. (ed.), 1984, pp. 168-76.

McCarthy, J. and Hayes, P., 1969
Some Philosophical Problems from the Standpoint of Artificial Intelligence *Machine Intelligence* 4, pp. 463-502.

Michie, D. (ed.), 1979
Expert Systems in the Micro-electronic Age (Edinburgh University Press).

Michie, D., 1982
Machine Intelligence and Related Topics (Gordon and Breach).

Minsky, M.L., 1975
A Framework for Representing Knowledge in Winston, P.H. (ed.), 1975, pp. 213-77.

Mylopoulos, J., 1980
An Overview of Knowledge Representation in Brodie, M.L. and Zillies, S.N. (eds), 1980, pp. 5-12.

Naylor, C.M., 1983
Build Your Own Expert System (Sigma Technical Press).

Newell, A., 1982
The Knowledge Level *Artificial Intelligence* 18(1), pp. 87-127.

O'Shea, T. and Eisenstadt, M. (eds), 1984
Artificial Intelligence: Tools, Techniques and Applications. (Harper and Row).

Quinlan, J.R., 1982
Inferno: A Cautious Approach to Uncertain Inference (Rand Corporation, California. N-1898-RC).

Reilly, P. (forthcoming)
Computers in Field Archaeology: Agents of Change? in Cooper, M.A. and Richards, J. (forthcoming) *Current Issues in Computer Archaeology* (British Archaeological Reports).

Salmon, M.J., 1982
Philosophy and Archaeology (Academic Press).

Samuel, A.L., 1963
Some Studies in Machine Learning using the game of Chequers in Feigenbaum, E.A. and Feldman, J. (eds). 1963, pp. 71-105.

Shortliffe, E.H., 1976
Computer-Based Medical Consultations: MYCIN (Elsevier, New York).

Sloman, A., 1971
Interactions between Philosophy and Artificial Intelligence: The Role of Intuition and Non-logical Reasoning in Intelligence *Proceedings of the Second Joint Conference on Artificial intelligence* pp. 270-78.

Sloman, A., 1979
Epistemology and Artificial Intelligence in Michie, D. (ed.),1979, pp. 235-41.

Sowa, J.F., 1984
Conceptual Structures: Information Processing in Mind and Machine (Addison-Wesley).

Stefik, M., Aikens, J., Balzer, R., Bendit, J., Birnbaum, L., Hayes-Roth, F. and Sacerdot, E., 1982
The Organisation of Expert Systems, A Tutorial *Artificial Intelligence* 18(2), pp. 135-73.

Szolovits, P. and Pauker, S.D., 1978
Categorical and Probabilistic Reasoning in Medical Diagnosis *Artificial Intelligence* 11, pp. 115-44.

Szolovits, P. (ed.), 1982a
Artificial Intelligence in Medicine (American Association for the Advancement of Science Selected Symposium 51, Westview Press, Colorado).

Szolovits, P., 1982b
Artificial Intelligence and Medicine in Szolovits, P. (ed.), 1982, pp. 1-19.

Torrance, S. (ed.), 1984
The Mind and the Machine: Philosophical Aspects of Artificial Intelligence (Ellis Horwood).

continued on p 12.

which display substantial and hence securely measurable abrasion on their corners. It must be emphasised that these measurements cannot be directly compared between brooches, but only in relation to each individual brooch.

These brooches were either worn singly or in pairs and occasionally as a triplet, and inspection of the results suggests a significant correlation between patterns of abrasion and their use in one of these fashions. Conclusions are here focused only on the measurements for top corners. We find that of the seventeen headplates of brooches discovered singly, no less than thirteen (76 per cent) are more abraded on the top *left* corner than on the top right (these larger abrasions are underlined in Table 1a for the upper and lower corners separately). Of the remainder, two show no significant difference, and only two are more abraded on their top right corners.

Of the eighteen pairs of brooches where measurements were possible on both top corners (see Table 1b), ten pairs (56 per cent) have *both* brooches more abraded on the top *right* corners than on the top left. Of the remainder, three pairs have one brooch abraded more on the left and the other more on the right; one pair has only one brooch showing a significant difference (again on the right); and of a further three pairs where one of the pair does not have measurable top corners, the other brooch again has its right top corner more heavily abraded. In only *one* case (the pair from Lyminge gr. 44) do both brooches have their top left corners more heavily worn.

In general, then, single brooches are more abraded on their top left corners; paired brooches are usually *both* more abraded on their top right corners. No statistics are required to see that these results are highly significant. But what do they signify? Clearly there is evidence here of the way in which brooches were usually worn and this shows, in the first place, that there was a consistency in orientation of brooches on costumes at this period which we might not have suspected. Furthermore, not only were single brooches generally worn in one orientation, but both members of a pair were often worn in another orientation. Figure 1c demonstrates the possible combinations of orientation and wear patterns for a pair of brooches abraded on the top corners, allowing for the manner in which brooches tend to hang forward from their fixing point, while Figure 1d shows possible orientations for a single brooch abraded on its top left corner. The dotted panel in Figure 1c shows the most common orientation implied by abrasion on top right corners. We may conclude that the single brooches were worn with their headplates pointing leftwards (since the top left corners are more commonly abraded) and pairs with their headplates pointing to the right, and probably parallel (both top right corners more commonly abraded).

Unfortunately the quality of excavation records for most of these brooches, many excavated in the last century, permits few reliable confirmations. The more recent account of the discovery of the Lyminge grave 44 pair does, however, perfectly conform to our conclusion for that aberrant pair (Warhurst 1955, 30f.) (one was lying 'at the pelvis, head pointing towards the right elbow' (i.e. to the left) and the other 'at the same angle, a few inches below it'). These Lyminge brooches are among the largest of the Kentish brooches, and in this respect are more like the single brooches. It may be that here lies some partial explanation for the unusual wear patterns on this pair.

The patterns of abrasion for lower corners show no such simple relationship, and may reflect the more complex and perhaps random factors of brooch and cloth usage which future work along these lines may perhaps be able to establish.

It is not possible or necessary here to follow the further implications of this discovery as they affect our view of dress fastenings and clothing fashions in the sixth century. The above example has served to demonstrate the usefulness of this approach to measurement of abrasion on jewellery. It has been concerned with just one class of brooch. The possibilities for circular brooches are clearly not quite so promising, but even here observations have already been made of uneven abrasion (e.g. Avent and Evison, 1982, p. 92). The same principles could be extended to other clothing accessories at this and other periods, particularly where there are features of their profile which are susceptible to differential wear.

References

Avent, R. and Evison, V.I., 1982

Anglo-Saxon button brooches *Archaeologia* 107, pp. 77-124.

Chadwick Hawkes, S. and Pollard, M., 1981

The gold bracteates from sixth-century Anglo-Saxon graves in Kent *Frühmittelalterliche Studien* 15, pp. 316-370.

Warhurst, A., 1955

The Jutish cemetery at Lyminge *Archaeologia Cantiana* 69, pp. 1-40.

Continued from p.7

Weiss, S.M. and Kulikowski, C.A., 1984

A Practical Guide to Designing Expert Systems (Chapman and Hall).

Weizenbaum, J., 1976

Computer Power and Human Reason: From Judgement to Calculation (Freeman).

Wilcock, J.D. (forthcoming)

A Review of Expert Systems, Their Shortcomings and their Possible Applications in Archaeology *Computer Applications in Archaeology* 1985.

Winston, P.H. (ed.), 1975

The Psychology of Computer Vision (McGraw-Hill).

Zadeh, L.A., 1979

A Theory of Approximate Reasoning *Machine Intelligence* 9, pp. 249-94.