NEW ZEALAND ARCHAEOLOGICAL ASSOCIATION

ARCHAEOLOGY IN NEW ZEALAND

BEYOND GIS: ARTIFICIAL INTELLIGENCE IN ARCHAEOLOGY

Claire Reeler
The University of Auckland

This paper introduces a new form of analysis, currently being used in the study of the internal organisation of pa sites. A new methodology, based on the use of artificial neural networks and fuzzy logic is presented, and a few of its applications into archaeology are discussed. The methodology is designed to enhance traditional forms of analysis using statistical methods and GIS (Geographic Information Systems). It introduces a further dimension into these and other forms of analysis.

The use of GIS in archaeology is becoming well known. Various software packages have been used to address questions in landscape and settlement archaeology (for example Kvamme 1989; Allen 1990; Carmichael 1990; Green 1990; Hasenstab and Resnick 1990; Madry and Crumley 1990; Brandt, Groenewoudt and Kvamme 1992; Hunt 1992), as well as intrasite analysis (Reeler 1992). GIS in archaeology is the focus of one of the largest colloquia at the forthcoming XIII Congress of the UISPP in Forli, Italy, in September this year. It is a growing field that is becoming well established.

The use of artificial neural networks and fuzzy logic represents a step beyond GIS into the realm of artificial intelligence. These are both techniques that are important tools in the study of artificial intelligence in the fields of engineering and computer science, as well as other disciplines such as psychology and economics. Much software and hardware is being developed which is opening up a new range of possibilities for those interested in modelling the way that humans think and learn. These developments have enormous potential advantages for researchers in all disciplines. They have particular relevance to Anthropology and Archaeology because of their potential to allow us to analyse the way in which humans view the world, and to simulate human thought processes. Neural networks and fuzzy logic are separate subdisciplines of artificial intelligence, but can be used together very effectively.

At the simplest level, GIS provides tools for organising data and asking questions such as 'where'. Similarly statistical techniques extract information from data based on questions such as 'how many', or 'how often'. Neural networks and fuzzy logic allow us to address questions of 'why' and to extract 'rules' from patterns. Artificial neural networks are essential designed not only to be good at pattern recognition, but to approach the recognition of patterns in the same way that a human being would approach them. It may be argued that many forms of archaeological analysis are directly involved in pattern recognition, and artificial neural networks are therefore ideally suited to archaeological analysis.

ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are computer simulations of the biological brain in its most basic form. The capacity of biological brains to learn and recognise patterns are two of the main foci of the simulations. Artificial neural networks are based on the way that neurons interact in biological brains. Programmers make certain assumptions about the way that a brain works when they create neural networks. These assumptions cover both the training of these networks, and the way that they analyse information. In a biological brain neurons are connected by axons, along which information is passed. When information is passed from a neuron along an axon, the neuron is said to have 'fired'. Different firing patterns are thought to relate to specific items of 'knowledge'. The designers of neural networks claim that when one learns something one is basically learning to fire neurons in a specific pattern. Neural networks are designed to 'learn' by recognising patterns within data (Aleksander and Morton 1990).

There are two main types of artificial neural networks - those that learn to classify data based on previously encountered data, and those that seek patterning in data independently to any form of classification. Both types of artificial neural networks can be readily applied to archaeological problems, and will be briefly described.

Artificial neural networks that classify are exposed to training data, which is a collection of information similar to that which will be analysed. It is very important that this data is truly representative of the data that the artificial neural network will be asked to analyse, but it does not need to be (and in fact should preferably not be) the same as that used for the analysis. A small subset of the full data for analysis is usually used as training data, but it is important that this small subset typifies the data as a whole. Artificial neural networks are repeatedly shown this training data and learn to associate patterns within the data with the classifications. When they are later shown real data to analyse, they look for the patterns and classify accordingly. This is a similar process to the way an archaeologist learns to classify sites after repeatedly encountering them in the field, or artefacts after repeatedly analysing them.

An important point to realise is that artificial neural networks do not require absolutely comprehensive data because they classify according to patterns in the data, and not just the data itself. In other words, if a site fulfills the criteria for a certain classification in most of its attributes, although not in all, the artificial neural network will classify it as 'most probably' of that classification. This is the same way that a person can recognise a category of artefact from only a few facts about that artefact, as long as those facts fit some pattern that the person has learnt. The process of learning and the resulting knowledge, which we term 'experience', is what is being modelled by artificial neural networks.

Non-classifying artificial neural networks merely look for patterns inherent within the data, and produce maps of this patterning. This can be useful in cases where one wishes to do analysis that is independent of set classifications, or even when one wants to test whether or not the patterns within the data correspond to those which one is using for classification.

FUZZY LOGIC

Fuzzy logic is based on the work of Zadeh (1965) who described the mathematics of fuzzy set theory and thereby the precepts of fuzzy logic. The idea behind 'fuzzy' concepts is that many real world ideas are vague or imprecise (Kosko 1993). Conventional Boolean logic relies on the fact that things are either one thing or another. It can therefore readily be rendered into a binary format. This binary format is very useful for computer analysis. However, in everyday language we use many concepts that cannot easily be translated into mutually exclusive categories. Some of the most commonly used examples are those of ideas such as 'tallness' (Brule 1995).

When we describe someone as 'tall', we do not have a numerical height in mind. Whilst there may be a range of heights that would be incorporated into the notion of 'tall', the limits of these ranges would be very difficult to define. Furthermore, these ranges would be culturally dependent on the ranges of heights within given populations. The difficulty in defining these ranges in specific ('crisp') terms can be illustrated by the example that if one asks several people to define the absolute lower numerical limit of the concept 'tall', one gets answers that vary between 1.5 and 1.8 m. (Most people in fact have a great deal of difficulty in assigning numbers to their own concept of 'tall' and one of the most common definitions seems to be that 'tall' is anyone whose height is greater than one's own. This illustrates the point that humans cannot easily resolve all problems into a binary format). Furthermore, if one were to accept any one of those heights as the lower limit of 'tall', for example 1.6 m, it would be hard to accept in real terms that a person who is 1.61 m is 'tall' and one who is 1.59 m is 'not tall'.

Fuzzy systems (using fuzzy logic) approach this problem by assigning memberships to a concept, within the range 0.0 to 1.0. For example, a person with a height of 2 m would belong to the category 'tall' with a membership value of 1.0. A person with a height of 1.8 m would belong to the category 'tall' with a membership value of 0.6 and a person with a height of 1.6 m would belong to the category 'tall' with a membership value of 0.2. The person with a height of 1.6 m might also belong to the category 'medium height', with a membership value of 1.0. Although there are certain basic guidelines available with respect to assigning membership values, it is essentially up to the researcher to decide how to apply the range of possible membership values to their data. The intuitive decisions about whether 1.6 m is 'tall' or not, still need to be made. Fuzzy systems merely help these decisions to be made by allowing one to use a range of values, instead of specific values. It is important to note that memberships are not mutually exclusive. They appear similar to 'probabilities' in mathematics, but as are illustrated by the above examples, are not probabilities, since they are not bound by having to add up to 1.0. In fact, probability theory is subsumed within fuzzy systems theory, and fuzzy systems themselves are subsumed within 'possibility' theory (Brule 1995).

Fuzzy logic includes the use of adjectives such as 'very' or 'somewhat'. These are termed 'hedges' and can also be defined mathematically. It has been argued that any idea that can be phrased in language can be rendered into fuzzy systems (Brule 1985). The use of fuzzy systems therefore enables one to move between language and mathematical terms in a consistent way. This has important implications for computer analysis, which is based in mathematical terms. 'Fuzzy' concepts are ideally suited to the training of neural nets, since the concepts that our own brains learn are usually 'fuzzy' in nature. Similarly, fuzzy systems are ideally suited to an analysis of human behaviour, since none of us would deny that the ways in which humans behave are seldom a result of easily defined opposing factors. Part of the human experience is the awareness that reasons for behaviour are seldom 'black' and 'white', but more often 'shades of grey' - to use popular idiom. It is therefore important when attempting to model and understand human behaviour, to have a method that allows one to build fuzziness into the analysis.

Artificial neural networks can analyse fuzzy data as easily as binary data, given the correct algorithms, and can therefore identify 'fuzzy' patterns. The mathematical expressions of fuzzy logic become important when an artificial neural network is reporting that a particular piece of example data is 'mostly' like one particular classification, but also 'fairly' like another, and 'very slightly' like a third. Fuzzy logic provides a mathematical means to work with those degrees of certainty.

Fuzzy logic also has another important advantage when used in the application of artificial neural networks. It can be used with classifying artificial neural networks to extract the 'rules' which the artificial neural network is using to decide on the classification of the data. These rules are the rules that the artificial neural network has determined itself from exposure to the training data. They are the rules that the artificial neural network has decided are the basis of each classification, in terms of the training data it was given. By extracting these rules we can see the relative importance of various aspects of the data in terms of the classifications.

Furthermore, since the artificial neural networks are designed to model the way that human beings learn and interpret patterns, we might hope that the rules used by the artificial neural network might model the sorts of rules used by humans as well. The potential for this sort of result to increase our understanding of past behaviour is enormous. Since the role of the archaeologist is to understand how humans thought and behaved in the past, being able to model the human brain in a way that reveals human perceptions has much potential. It might even be suggested that the use of artificial neural networks has the potential to transcend "culture" since the networks model the level of biology. It is implicitly assumed within the use of neural networks that they reflect the way that all humans will learn, despite their cultural background. By training a neural network one is therefore hoping to reproduce the way that humans learn to view the world. In fact, the influence of culture would be to apply specific patterns within the neural network, thereby producing a learned response. By extracting the rules that govern the patterning, one could be said to be extracting the cultural rules. It is hoped that the neural networks will find patterns inherent within the archaeological data that reflect the assumptions made by the people who produced the archaeological material. The use of fuzzy systems acknowledges the fact that these cultural rules will probably not be simple or clear cut. They will probably not be binary 'either... or' statements. They will probably require the use of adjectives such as 'very' or 'somewhat'.

Artificial neural networks and fuzzy logic can add an additional dimension to many forms of archaeological analyses. Typological analyses, whether of artefacts or sites, can use artificial neural networks and fuzzy logic to extract the rules for different classifications, as well as to identify classifications. Classifications into different categories themselves could gain an additional dimension by the use of fuzzy logic. For example, a site may be described as situated within a 'good' area for growing crops, 'fairly near' to a wide range of resources such as fresh water and various types of food and 'close' to the coast. The adjectives in inverted commas are all fuzzy and would be described by ranges of values. A 'good' area for crops would reflect certain ranges of values for soil type, drainage, rainfall, temperature and other factors. Activity areas might be described as 'primarily' concerned with food preparation, but 'slightly' involved with food storage, artefact manufacture and flax working as well. Again these would be defined by ranges of values representing the types of artefacts and features found within the activity area. Artificial neural networks and fuzzy logic could be applied to many of the analyses done with GIS in archaeology.

Artificial neural networks and fuzzy systems are currently being applied to a study of the internal structure of pa sites. This analysis will test the application of these methods into one form of archaeological analysis. The results of the analysis will be published at a later date. It is hoped that this description has introduced the idea of the application of these methods into archaeology, and illustrated their enormous potential within this discipline.

Personal details: I am currently enrolled in PhD in Archaeology at Auckland University, studying the internal organisation of pa sites. As a part of this study I have developed a database of information extracted from the published reports of pa that have been excavated. Several people have provided valuable help in this regard, amongst them Matt Schmidt and Tony Walton. Information from this database will be generally available towards the end of this study. My previous training in archaeology was at the University of Cape Town in South Africa, where I specialised in the spatial analysis of Late Stone Age coastal midden sites, using GIS.

REFERENCES

Aleksander I. & Morton H. 1990. *An Introduction to Neural Computing*. London. Chapman and Hall.

Allen K.M.S. 1990. Modelling early historic trade in the eastern Great Lakes using geographic information systems. In: Allen K.M.S., Green S.W. & Zubrow E.B.W. (Eds) *Interpreting Space: GS and Archaeology:* 319-329. London. Taylor and Francis.

Brandt R., Groenewoudt B.J. & Kvamme K.L. 1992. An experiment in archaeological site location: modelling in the Netherlands using GIS techniques. *World Archaeology* 24(2):268-282

Brule J. 1995. Fuzzy Systems - a Tutorial. Internet UseNet News comp.ai. May 1995.

Carmichael D.L. 1990. GIS predictive modelling of prehistoric site distributions in central Montana. In: Allen K.M.S., Green S.W. & Zubrow E.B.W. (Eds) *Interpreting Space: GS and Archaeology:* 216-225. London. Taylor and Francis.

Green S.W. 1990. Sorting out settlement in southeastern Ireland: landscape and geographic information systems. In: Allen K.M.S., Green S.W. & Zubrow E.B.W. (Eds) *Interpreting Space: GS and Archaeology*: 356-363. London. Taylor and Francis.

Hasenstab R.J. & Resnick B. 1990. GIS in historical predictive modelling: the

Fort Drum project. In: Allen K.M.S., Green S.W. & Zubrow E.B.W. (Eds) *Interpreting Space: GS and Archaeology:* 284-306. London. Taylor and Francis.

Hunt E.D. 1992. Upgrading site-catchment analyses with the use of GIS: investigating the settlement patterns of horticulturalists. *World Archaeology* 24(2):283-308

Kosko B. 1993. *Fuzzy Thinking*. London. Harper Collins.

Kvamme K.L. 1989. Geographic Information Systems in Regional Archgaeological Research and Data Management. In: Schiffer M. (Ed) *Archaeological Method and Theory.* 1: 139-203.

Madry S.L.H. & Crumley C.L. 1990. An application of remote sensing and GIS in a regional archaeological settlement pattern analysis: the Arroux River valey, Burgundy, France. In: Allen K.M.S., Green S.W. & Zubrow E.B.W. (Eds) *Interpreting Space: GS and Archaeology*: 364-380. London. Taylor and Francis.

Reeler C. 1992. Spatial Patterns and Behaviour at Dunefield Midden. Unpublished MA thesis, University of Cape Town, South Africa.

Zadeh L.A. 1965. Fuzzy Sets. *Information and Control* 8:338-353