24

Back-propagation algorithms to compute similarity relationships among archaeological artefacts

Juan A. Barcelo

(Universitat Autonoma de Barcelona, Dpt. Historia de les Societats Precapitalistes i Antropologia Social, Barcelona, Catalunya)

24.1. Introduction

Most archaeologists know the importance of a proper study of similarity relationships among archaeological artefacts. However, there has been some controversy describing the kind of inference derived from the detection of those statistical relationships. That controversy does not come from the nature of similarity, but from its way of calculation. Most archaeologists coincide in definitions like:

> "SIMILITUDE (or similarity): Formal resemblance in general details, but not in everything".

However, what does it mean *resemblance in general details*? Some researchers have proposed that this property can be measured selecting a reduced set of descriptive features. Those features are selected because of their *relevance* in an explicit task. In this context, *relevant feature* is used as a synonym of *intrinsic property*. Adams (1988) defines the term "intrinsic property" as a property of an object that is inherent to the object itself, and is independent of the properties of other objects or the deposit with which the object is associated.

The opposite approach rejects any selection of descriptive features based on their contextual or intrinsic relevance. It adopts instead a more "universal" method based on the quantitative evaluation of the number of common features. In this context, *similarity* is a mathematical function measuring the number of commonalties among a series of archaeological artefacts.

The Spaulding-Hodson debate (see Whallon and Brown 1982) exemplifies this controversy. Are types really a product of the grouping of objects or are a result of the variables (or attributes) association? Hodson asserted that seeking associations between attributes is quite different from object clustering, and that an archaeological *concept* has to be defined in terms of the objects explained by that concept (Hodson 1980). Spaulding argued, on the other hand, that there is no reason to work with "sets of objects"; the pattern of associations among the variables is a more pragmatic and easy way to understand archaeological types, specially if variables are defined in a nominal scale.

Cowgill (1990) and Read (1989) try to answer to this dilemma arguing that both positions are only two different levels of analysis. Cowgill thinks that the debate has been misguided because some researchers have been talking about searching for patterning *within* a single assemblage or among a group of very closely related assemblages while others have intended comparisons *between* or *across* a number of different assemblages. That is, archaeologists do not recognise whether types (or archaeological concepts) are discovered or imposed because there is not any coherent discussion about the *level* of analysis. The possibility of making worthwhile inferences about categories meaningful to the ancient makers and users depends on the making clear whether one is talking about classes generated by considering objects from a single assemblage (or at most a group of very closely related assemblages) or classes generated by considering objects from diverse assemblages. "I suspect that most archaeologists are so deeply committed to using classification for between-assemblage studies that they cannot assimilate the idea of classifications dedicated to the search for patterning within a single assemblage" (Cowgill 1990: 67).

Read (1989) asserts that *morphological similarity* is the consequence of clustering, and not its cause. It is a good inference mechanism to measure the pattern of associations *within* an assemblage, but the archaeologist has to build that assemblage before he/she computes similarity algorithms. "If structuring processes are the beginning point of understanding the data in hand, then the initial goal becomes one of relating structuring process to measurable groups in the data and not the reverse" (Read 1989, 184).

In this paper I will try to show how this debate has been affected by the statistical methods used by archaeologists. Neural networks and connective algorithms allow archaeologists to discover the importance of *object groupings* by accepting also the relevance of the pattern of associations among the variables: groups of artefacts are not really useful for an archaeologist if they are not expressed in a specific formal language. Thus, the goal of pattern of associations is to *represent* some archaeological concept. We can use a neural network to translate a group of objects into a pattern of associations.

24.2. Concepts or types?

Why should we discover, describe and explain "similarity"? I think that by analysing similarity relationships we not only build groups, but we define *concepts* or *theoretical entities* as well. The detection of similarity is a method to build archaeological concepts; probably it is not the only inference method we have to carry on this task, but one of the most important tools in archaeological practice.

What is a *concept*? Psychologists say that concepts reflect the way that we divide the world into classes, and much of what we learn, communicate, and reason about involves relations among these classes (Sternberg and Smith 1988).

Concepts are our means of linking perceptual and non-perceptual information. For example, we use a perceptual description of the grave goods in a burial and then use our non-perceptual beliefs about the social nature of material culture to direct our interpretation: "a grave with a sword among the grave goods belongs to a high social category". Concepts are recognition devices because they serve as entry points into our knowledge stores and provide us with expectations that we can use to guide our actions. By partitioning the archaeological record into classes using explanatory concepts, we decrease the amount of information we must perceive, learn, remember, communicate and reason about.

It is important that we see archaeological types as *concepts* and not as pure sets of related artefacts or related variables. They are the *theoretical entities* we need to organise archaeological knowledge into archaeological theories. In this paper I do not deal with the classical debate about the nature of theoretical entities (see Tuomela 1973, Niiniluoto and Tuomela 1973, Rivadulla 1986, Hooker 1987, Churchland 1988, Thagard 1988, Gooding 1990), although it has a great importance to identify concept formation mechanisms. In essence, the controversy about the real nature of concepts comes from the fact that they are mental representations, and not empirical categories. This fact is also true about archaeological types, that is, they are virtual entities without existence out of the archaeological theory they contribute to organise. Consequently archaeological concepts are unobservable. We have access to them only in terms of their instances.

To have a concept of *X* is to know something about the properties of *X* 's instances. Most archaeologists seem to think that the properties contained in a concept are *singly necessary* and *jointly sufficient* to define that concept (Adams 1988). For a property to be singly necessary, every instance of the concept must have that property; for a set of properties to be jointly sufficient, every entity having that set must be an instance of the concept. Such properties are referred to as *defining* or *intrinsic* for collectively they constitute a definition (Shoemaker 1980). According to this position, an archaeological artefact will be categorised as an instance of a concept if and only if it contains the defining properties of the concept.

The greatest shortcoming of the aforementioned position is the "practical" failure in discovering true definitions of concepts. Archaeologists are incapable of specifying the defining (or "intrinsic" in Adams' sense) properties of even the most simple concept. Different archaeologists use different definitions for the same concept. Even worse, they are not able to realise that they work with different entities. The reader may think that this is only a consequence of the failure of positivism, but the consequence is the same: there is not any single way to distinguish between intrinsic properties and descriptive features. Of course, we can use "subjective" evaluations about what is intrinsic or relevant and what is pure descriptive however, I believe that we cannot find any profitable solution in this way.

The non perceptual nature of concepts prevents the definition of their intrinsic properties. Is there any way to solve this epistemological problem? Psychologists use the notion of "prototype" to study the nature of concepts (Rosch and Mervis 1975). If there is no mechanism that enables us to decide which are the relevant intrinsic properties to define a concept, then we have to study the *most typical instance* of the concept to learn something about its nature. In other words, the definition properties of a concept are assumed to occur only in some instances, which contribute to define a *prototype*, for it describes accurately only the "best examples" of a concept. Therefore, the content of a concept is its prototype, and an object will be categorised as an instance of a concept if it is sufficiently similar to the prototype, similarity being determined, in part by the number of properties that the object and the prototype share.

How can archaeologists use the notion of *prototype*? Methods used by psychologists to build concepts from prototypes are not useful to us. Psychological research is based on interviewing volunteers; given a concept, people are asked about the degree of typicality of some of its instances; the arithmetic mean of all subjective evaluations characterises the concept. One way to build archaeological concepts may be this: look for consensus. However, I think that archaeologists need more reliable methods; therefore, I propose the use of *experimentation.*

An experiment is nothing more than a series of *controlled* observations. For example, if we want to *experiment* with the chronology of a particular pottery shape, we must observe a data set (or *control* set) which includes only instances of the concept, e.g. "20th Century". This controlled observation allows us to transfer the descriptive information contained in the instances, into the representation of the concept. *Induction* may be defined as the *transference* of information from a set of data items (observed in the empirical level) to the *representation of a concept* (Holland *et al.* 1986). The "representation of a concept" is a physical entity only because we can build it as a computational unit able to store the *induced* knowledge.

Learning, *generalisation* or *induction* (different names for the same reasoning mechanism) is a phenomenon exhibited when a system improves its performance at a given task without being reprogrammed (Forsyth 1989); in other words, not every transference of knowledge from data to concepts is profitable. Only if the experimental induction leads us to the improvement of a previous concept, then we will accept the new definition. Improvement can mean various things, including: a higher proportion of correct decisions, faster response, lower-cost solutions and wider range of applicability. In all of these cases, *induction* implies a comparison with another situation or context: we cannot learn anything unless its performance can be asserted reliably. For induction to take place, there must be a criterion according to which decisions or partial solutions can be scored. Otherwise no one can say whether the system has changed for the better or for the worse. Then, evaluative feedback is absolutely fundamental to the archaeological concept formation.

To improve a representation we should define more control sets. For example, one with instances of the concept "21st Century", and the other with instances of the concept "19th Century". For pragmatic reasons, the best representation of our concept will be the most discriminative one. It is easy to understand that the typical variations from typicality observed during experiments or controlled situations are our source information to build prototypes; an instance is typical to the extent that its properties occur frequently in the concept.

There is nothing really new in this "experimental" approach to prototyping. In archaeology, Cowgill (1990) has asserted that the analysis of patterning within a single assemblage must precede the analysis of patterning among a series of empirically unrelated assemblages. We have only to insist in the *controlled* nature of the assemblage: not all data sets are appropriate for prototyping, because we need to know before that all items in the assemblage are instances of a specific concept. Psychologists call this inference operation *supervised learning* (Fisher and Pazzani 1991). In early psychological studies, subjects were assumed to have discovered the appropriate concept after correctly predicting class membership for a sufficient number of consecutive trials. These studies assumed that the experimenter classified possible instances as "positive" or "negative", according to what he/she previously knew about the concept.

Concepts are used to explain empirical data. If and only if there is a relationship between the empirical data and the concept, the archaeologist will explain the archaeological record. Therefore, "using a concept" is equivalent to defining the relationship between a concept and its instance (Thagard 1989, Barcelo 1993a). However, we have built the representation of the concept according to what we know about some of its instances. It is not exactly true that an archaeological artefact is an instance of a member if it is sufficiently similar to known concept instances. The *controlled* observation or experiment needed to define a prototype does not operate on the basis of content-free general inference rules but, rather, is often tied to particular bodies of knowledge and is greatly influenced by the context in which it occurs.

Concepts are taken to be mini-theories about the nature of the categories they describe because we have integrated the experiment, the context of use, and the prototype into a single entity to be able to define that concept. Conceptualising an archaeological artefact is then a matter of applying the relevant theory, and consequently the definition of the concept depends on its explanatory power. The way we use concepts affects the way we can represent them.

The proposal that a concept includes only a prototype, turns out to be too simple because it is hard to see how a prototype can integrate all the knowledge needed in concept formation. Using prototypes to define concepts can be misguiding because of the potential variability among the instances: the concept definition varies each time we compute an instance, because we are introducing more variability into the concept representation. Psychologists think that concepts contain at least two components: a prototype and a *core* (Michalski 1989). The cores of fuzzy concepts are not definitions in the usual sense, but *meta-knowledge* units. In other words, we are speaking about knowledge on the way to induce the concept (Pitrat 1989). The inevitable result of this fact is that any representation of a concept can be never complete, because it is related with a fixed context of use. There are always many possible ways to use a concept in a theory.

Concept representations are intrinsically unstable, because they depend on the meta-knowledge available in each time. Cores do not need to be fixed, because scientists may change their theories and the way they plan to use concepts. However, not only concept representations vary across the different contexts of use, the properties and the pattern of associations among all instances also vary (Barsalou 1989). Instability both between and within individuals for graded structure may reflect uncertainty about the status of atypical instances. Because archaeologists lack knowledge for these exemplars or are unsure about their category membership, they may frequently change their minds about these exemplars' typicality and thereby produce instability. According to this view, stability should increase monotonically with typicality.

As we have seen, prototype properties tend to be perceptually salient, although not perfectly diagnostic of concept membership; in contrast, the properties that comprise the core are most diagnostic of concept membership but tend to be relatively hidden. Because the core properties are not easily accessible, we have to use heuristic rules to decide the experimental context in which we will calculate prototypicality. We confront some representation of an instance with our knowledge of the various categories it might belong to. The instance in one of these categories will produce a reasonable explanation of the information we have about it, if and only if this explanation is better than the one produced by other candidate categories. Only in those circumstances we will infer that an instance is a member of a category (Murphy and Medin 1985, Wattenmaker et al. 1988, Rips 1989).

24.3. Advances in the mathematical study of similarity relationships

One of the main strategies to define *archaeological concepts* has been to sort a specific collection of objects into groups, which members (in terms of properties salient to the analyst) tend to be much like one another and considerably unlike objects assigned to other groups, and then trying to see what distinguishes the groups from one another. Cowgill (1990) has characterised this approach by the phrase: "internal cohesion and external isolation".

We have many different mathematical techniques that are powerful for separating and/or aggregating data viewed as points in some *n* dimensional space. But are the premises of these techniques the premises under which culturally based distinctions were made? Classical approaches to similarity relationships are based on the assumption that perceived *dissimilarity* is associated with conceptual distance (or typicality). This approach represents the observations as points in a multidimensional metric space, and assumes that judgments of the perceived similarity of two stimuli are inversely related to the distance between their perceptual representations (Shepard 1987, Sokal 1988). This class of models, known as the *geometric model of similarity*, is contained within the larger class of multidimensional scaling (MDS) models. MDS models do not necessarily require the perceptual space to be metric. If we focus on the simplest case (two-dimensional space and Euclidean metric), the MDS model assumes that:

$$d\,(a,\,b) = [(x_a - x_b)^2 + (y_a - y_b)^2]^{1/2}$$

where $d(a,b)$ is the calculated *dissimilarity* , x_a is the coordinate of observation a on dimension x of the conceptual space.

Because of their reliance on distance, geometric models predict that perceived dissimilarities must satisfy certain distance axioms:

- *the self-dissimilarities of all observations are equal.* There are some problems with this axiom: distinctive or unique archaeological artefacts, or objects having few features in common with other objects have a greater perceived self-similarity, and so a smaller perceived self-dissimilarity (see Krumhansl 1978).
- *Minimality.* Two different artefacts are always at least as dissimilar as either artefact is to itself.
- *Symmetry.* The asymmetry underlying most artefact matching shows that this axiom is also inappropriate in archaeological research (see Tversky 1977, Glucksberg and Keysar 1990, on the asymmetry of similarity relationships).
- *Triangle inequality* (or Ultrametric Axiom). A century ago, William James (1890) gave an example of what seem a clear violation: a flame is similar to the moon because they both appear luminous, and the moon is similar to a ball because they are both round. In contradiction to the triangle inequality, a flame and a ball are very dissimilar. This kind of violation is very usual in ethno-archaeological research (Barcelo 1993b).

We have just seen that a concept or theoretical entity is more than a detected correlation among some descriptive features. Psychologists assume that comparison statements of the form *a is similar to b* are assessed by comparing features of a with features of b. Obviously not all the features of a and b are considered; instead, only a relevant subset of the features of a and b is selected before any comparison or matching operation. The number of features that can be attributed to any given object is unlimited; any theory of feature matching must postulate prior feature selection. We have seen that this *selection* is a conceptual operation affected by the context in which it is produced and the specific goals, which underlie the matching.

Nevertheless, traditional geometric similarity models are context free: they predict that the similarity between a and b is only a function of the distance between their descriptions and so is unaffected by how many or which other stimuli (different instances or different concepts) are in the domain. We have just seen that this assumption is wrong. Any two things picked at random *must* be similar to one another in at least some respects, because the particular ways that any two things resemble are always determined by the context. Also because of their questionable empirical validity, it is desirable to investigate theories of perceived similarity not constrained by the distance axioms.

Some authors have introduced a weighted Euclidean model, more akin with the new theoretical discussion of the nature of concepts:

$$d\,(a,\,b) = [w^2_{x,j}\,(x_a - x_b)^2 + w^2_{y,j}\,(y_a - y_b)^2]^{1/2}$$

where w_i $(i = x, y)$ is a weight reflecting the importance that subject j places on dimension i. Alternatively, the weights may be interpreted as measures of relative selective attention. Under this interpretation, each weight measures the degree to which an archaeologist attends to a dimension (see other modifications of the same model in Ashby and Perrin 1988).

One of the most influential approaches to solve the shortcomings in the classical model has been Tversky's contrast model (Tversky 1977). This model postulates prior extraction of those features that are relevant to the task; thus, the representation of an object as a collection of features is viewed as a product of a prior process of extraction and compilation. Once a relevant subset of features has been selected, the perceived similarity, s, between two objects, a and b, is considered to be a weighted function of selected features that are both common and distinctive:

$$s(a,b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

where θ reflects the weight assigned to features common to objects a and b, α the weight assigned to features of a that are not included in b, and β the weight assigned to features of b that are not included in a. Consequently, the similarity of an object a to an object b is conceived as a "linear contrast" of three, weighted feature sets: the features common to a and b, the features distinctive to a compared with b, and the features distinctive to b compared with a.

Weighting is a way to introduce the influence of context in the similarity calculation. Weights measure the *salience* of a feature, that is, their informational value. Some authors (see Ortony 1979, Osherson 1987, Glucksberg and Keysar 1990) have enhanced the model to have a more appropriate definition of this function: the salience, or weight, of the matching properties in a, is made to be dependent on the salience value of the matching properties in b. The resulting *weighted* similarity relationships have some interesting properties, among them the asymmetry in similarity relations: the recognised relationship between a and b differs from the recognised of b to a. Such asymmetries are the result to the differential salience of the features of a and b, which is another way to explain the context influence on similarity.

Modern accounts of this problem use a structure known as *confusion matrix* (Fig. 24.1). Each row in a confusion matrix is associated with one controlled observation, and each column with the concept to which this observation is related. The entry in row i and column j is an estimate of $P(Rj/Si)$, *or* the typicality measure associated with each concept. In this case, we are introducing not only *salience*, but also the subjective effects derived from the relationships among concepts in a competitive network.

		CONCEPT *i*	CONCEPT *j*	CONCEPT *m*
EXPERIMENT A	Observation 1	X_{1i}	X_{1j}	X_{1m}
	Observation 2	X_{2i}	X_{2j}	X_{2m}
	Observation 3	X_{3i}	X_{3j}	X_{3m}
EXPERIMENT B	Observation 4	X_{4i}	X_{4j}	X_{4m}
	Observation 5	X_{5i}	X_{5j}	X_{5m}
	Observation 6	X_{6i}	X_{6j}	X_{6m}
EXPERIMENT C	Observation 7	X_{7i}	X_{7j}	X_{7m}
	Observation 8	X_{8i}	X_{8j}	X_{8m}
	Observation 9	X_{9i}	X_{9j}	X_{9m}

Figure 24.1: A confusion matrix.

There are different estimations of $P(R_j/S_i)$ (see Ashby and Perrin 1988, Nosofsky 1992), although most of them are based on the *biased-choice model*:

$$P(R_j / S_i) = \beta_j \eta_{ij} / \Sigma \beta_m \eta_{im}$$

Here $P(R_j/S_i)$ is a function of the similarity between the description of artefact S_i and the description of artefact S_j, denoted η_{ij}, and of the bias toward concept R_j, denoted β_j. This expression gives the "strength" of making a Category *j* response given presentation of stimulus (an artefact description) *i*; the final categorisation probability is determined by the ratio of this strength and the sum of strengths for all concepts in the network.

When applying this model, it is assumed that similarity is symmetric, and that all self-similarities are equal. I have already questioned the empirical validity of these assumptions. Yet despite their possible inaccuracy, the biased-choice model has been successful at predicting the results of many recognition experiments (although not any archaeological experiment).

As a result of psychological studies with the choice-biased and related models of *confusability*, it has been assessed that similarity and confusability co-vary. However, this fact does not imply that similarity is the only factor affecting confusability among different possible concepts. Prior knowledge on each concept included in the confusion matrix also explains why archaeologists often go wrong when when deriving instances and prototyping: an archaeologist with a lot of knowledge on concept R_A and very few on concept R_B will almost never confuse the most typical instances of the first with the most typical instances of the second. $P(R_B/S_A)$ will be very low.

Ashby and Perrin (1988) believe that, in the absence of prior knowledge bias, perceived similarity is proportional to the probability of a confusion. Thus, the perceived similarity between two artefacts *a* and *b* is naturally defined as the proportion of descriptive features associated with *a* falling in the conceptual space assigned to R_b in an unbiased two-choice recognition task. This is a clear improvement of traditional geometric similarity models, which are context-free; here, similarity is measured in terms of the probability of inter-experiment-confusion errors (Nosofsky 1992).

To add a new artefact or concept (or more prior knowledge) causes the archaeologist to readjust the amount of attention focused on each dimension, thereby stretching effectively some dimensions and shrinking others, and in so doing, changing the predicted similarities between all artefact pairs. Ashby and Perrin's model predicts that similarity is determined by the overlap of *perceptual* (descriptive) distributions; it is only under certain, very special conditions, that overlap and distance measures agree, and thus the general Euclidean scaling model is contained within the general recognition theory as a special case.

We may translate psychological research (see Ashby and Lee 1991, Nosofsky 1992) into archaeological terms, if we consider each archaeological artefact as an experiment, and if the features selected to describe this experiment have some connection with the categorisation (or conceptualisation) process. This effect can be represented as a point in a multidimensional space, where noise produced by descriptive bias is the cause of variability over experiments. Thus, we are assuming that a distribution of experiments (or controlled observations) is the appropriate descriptive representation of an archaeological concept (the weighted sum of all experiments is converted into a *prototype*). The archaeologist is assumed to divide the experimental space (the set of concepts defined into a specific archaeological theory) into conceptual regions. On each experiment, the archaeologist determines in which region the descriptive effect falls and then emits the associated concept. To understand the relationship between identification, conceptualisation, and similarity we have to understand the manner in which conceptual regions change during prototyping.

There is not any single way to account the effects of experimental context and prior knowledge into conceptualisation. Different statistical models have been designed to deal with different kinds of influence:

- the expected value of similarity is a function of the difference between the means of the distribution of psychological magnitudes and the variance-covariance matrix of the difference between psychological values (Ennis *et al.* 1988)
- the probability that an instance joins a concept is determined jointly by the current size of each concept (the number of instances we already know), the similarity of the instance to the concept's central tendency (current prototype), and the value of a "coupling" parameter, which is the free parameter in the model. Roughly, the probability that artefact *i* is classified as an instance of Concept *j* is found by

summing the similarity of i to each concept's central tendency, weighted by the category-label j probability associated with the concept (Anderson 1990)

- the relative frequency with which exemplar j is presented during experimentation and the similarity between that instance and the others determine the graded category structure. In other words, archaeologists make classification decisions on the basis of similarity and the *frequency* of experiments. It is important to realise, however, that similarity and frequency can exert mutual influence on one another, thereby making more intricate the relation between similarity, frequency, and categorisation (Nosofsky 1988)
- rather than group entities together on the basis of the similarity between two entities such as A and B, an archaeologist has to group entities together on the basis of the *conceptual cohesiveness* between A and B. The conceptual cohesiveness between two events depends not only on those events and surrounding events E (the set of controlled observations) but also on a set of concepts C that are available for describing A and B together, and that have been selected by the archaeologist as a representation of his/her goals. Thus, the conceptual cohesiveness between two events is a four argument function $f(A,B,E,C)$ in contrast to an ordinary similarity function of two arguments $f(A,B)$ (Stepp and Michalski 1986).

In this section we have studied the importance of the proper recognition of similarity relationships in prototyping: the more similar an instance is to the other members of its category and the less similar it is to members of contrast categories, the higher will be the typicality rating given to that instance (Nosofsky 1988). Only by studying similarities and differences among different instances of a concept during an experimental situation, we will be able to build reliable prototypes and useful approximations to the definition core of concepts.

Similarity judgments are based on people's representations of entities. Therefore, rather than beginning with certain mathematical procedures and asking what uses we can find for them, it would be better to begin with the problems we need to solve, and then to find the most convenient solution. Some archaeologists have also asserted the same point (cf. Read 1989): to be informative of a cultural system, a classification, that is, a structured set of concepts, must capture similitude and difference relevant to the meanings provided through the cultural system, and that is quite difficult using standard statistical techniques, as correlation or variance-based measures, because these algorithms are insensitive to the goals proposed by the archaeologists.

24.4. A back-propagation algorithm to compute similarity relationships

The purpose of last section has been to show some of the problems we have to solve in order to include context-sensitivity in similarity calculations. Up to now there is not any single algorithm able to solve all these problems. We are obliged to use one algorithm for every different task instead.

In this paper I present the back-propagation algorithm, first designed by Rumelhart, Hinton and Williams (1986). It is not the best of all prototyping algorithms, nor the answer to all theoretical problems we have studied so far, but it is useful in some specific tasks. An additional advantage is its availability as commercial and shareware software. Source code appears in Pao (1989). Present description of the algorithm is based on the vector notation by Stone (1986).

A back-propagation model is a special kind of Neural Network. There are eight major components in any neural network:

- a set of processing units
- a state of activation
- an output function for each unit
- a pattern of connectivity among units
- a propagation rule for propagating patterns of activities through the network of connectivities
- an activation rule for combining the inputs impinging on a unit with the current state of that unit to produce a new level of activation for the unit
- a learning rule whereby patterns of connectivity are modified by experience
- an environment within which the system must operate

Figure 24.2 shows a way to represent neural networks in terms of a confusion matrix. There is a set of processing units (inputs and outputs) indicated by circles in the diagram. Each unit has an activation value, which is passed through a function to produce an output value. This output value is transmitted to other units in the system, through a set of unidirectional connections. There is associated with each connection a real number, usually called *weight* or strength of the connection, which determines the amount of effect that the first unit has on the second. The pattern of interconnections is not fixed, rather the weights can undergo modification while processing is underway as a function of experience.

Any neural network (or parallel activation model) begins with a set of processing units. These units represent particular conceptual entities such as artefacts, descriptive features or concepts. We can distinguish the input units — representing descriptive features — from output units — representing concepts or theoretical entities; we also need a third layer of *hidden* units to store the connection weights among input and output units.

In addition to the set of units we need a representation of the state of the network at time t. This is specified by a vector of N real numbers representing the pattern of activation over the set of processing units. In most cases we know the activation state of the input units. For instance, below shows variables which describe the way we see some artefacts.

	Length	*Height*	*Width*
Artefact-1	12.5	4.2	3.1
Artefact-2	10.2	5.3	2.2
Artefact-3	9.5	8.3	4.0

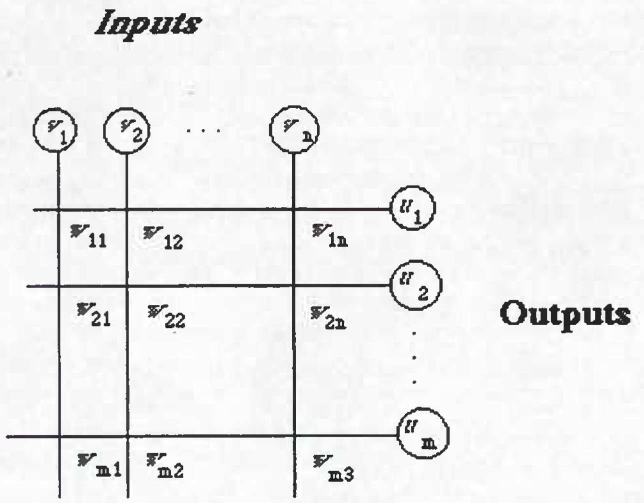


Figure 24.2: A matrix representation for neural networks.

Here we are working with a three input unit model (one for each variable), and we know three successive states of activation (one for each observation or artefact). The same can be seen for the output units (remember that we are in a *supervised learning* model). If our three observations have been rigorously controlled, we know that:

	Chronology
Artefact-1	11th Century
Artefact-2	10th Century
Artefact-3	9th Century

This network has only one output unit, and we know three successive states of activation for it. Units interact by transmitting signals to their neighbours, and therefore the degree to which they affect their neighbours, is determined by their degree of activation. Associated with each unit, u_i there is an output function $f(o_i)$ which maps the current state of activation $(a_i(t))$ to an output signal. In the back-propagation algorithm here explained, we assume the following sigmoid logistic non-linearity:

$$o_i = 1 / (1 + e^{-(net_j + \theta_j)/\theta_0})$$

The parameter θ_j serves as a threshold or bias (*learning rate*). Each unit provides an additive contribution to the input of the units to which it is connected. In such cases, the total input of the unit is simply the weighted sum of the separate inputs from each of the individual units. That is, the inputs from all the incoming units are simply multiplied by a weight, and summed to get the input to that unit.

$$net_j = \Sigma w_{ij}\, o_i$$

Given the fact that each unit is connected to all other units in the network, a positive weight represents an excitatory input and a negative weight represents an inhibitory input. We use here a weight matrix W to represent such a pattern of connectivity. The weight w_{ij} is a positive number if unit u_j excites unit u_i, it is a negative number if unit u_j inhibits unit u_i; and it is 0 if there is no direct connection among the units. Units in each layer (input, *hidden* and output layers) are not connected among them.

Let us see how a back-propagation model operates:

Step 1.
Initialise weights and offsets (or learning rates). Set all weights and node offsets to small random values

Step 2.
Present input and desired outputs. Present a continuous valued input vector $(x_0, x_1, \ldots x_{n-1})$ representing one single observation (for example the description of an archaeological artefact), and specify the desired output $(d_0, d_1, \ldots d_{m-1})$.

Step 3.
Calculate actual outputs. Use the sigmoid non-linearity output function $f(o_i)$ and the following formulae to calculate outputs $(y_0, y_1, \ldots y_{m-1})$:

$$y_m = f(\Sigma w_{km} x_k - \theta_j)$$

$$x_k = f(\Sigma w_{jk} x_0 - \theta_\varphi)$$

Note that y_m is the set of units in the output layer. Their respective values have to be calculated from the values in the previous (*intermediate or hidden layer*) one.

Step 4.
Adapt weights. Use a recursive algorithm starting at the output nodes and working back to the hidden layer. Adjust weights by:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i$$

In this equation $w_{ij}(t)$ is the weight from hidden unit i from an input to unit j at time t. x_i is either the output of unit i or is an input, η is a scalar constant, which determines the rate of learning, and δ_j is an error term for unit j (it can be seen as the difference between the desired and actual output on input i). If unit j is in the output layer, then:

$$\delta_j = y_j (1 - y_j)(d_j - y_j)$$

where d_j is the desired output of unit j and y_j is the actual output produced in Step 3. If unit j is an internal hidden layer, then

$$\delta_j = x_j (1 - x_j) \Sigma \delta_k w_{jk}$$

where k is over all units in the layers above unit j. Internal unit thresholds are adapted in a similar manner by assuming they are connection weights on links from auxiliary constant-valued inputs. Convergence is sometimes faster if a momentum term is added and weight changes are smoothed by:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_j x_i + \alpha (w_{ij}(t) - w_{ij}(t-1))$$

where $0 < \alpha < 1$.

Step 5.
Repeat by going to step 2. Take another vector input representing the next empirical observation, and start the iterative processing again.

This algorithm is typically applied to the case in which pairs of patterns, consisting of an input pattern and a target output pattern, are to be associated so that when an input pattern is presented to an input layer of units, the appropriate output pattern will appear on the output layer of units. The procedure is an iterative gradient algorithm designed to minimise the mean square error between the calculated output and the observed one in an experimental situation.

To sum up, back-propagation is a supervised learning technique that compares the responses of the output units to the desired response, and readjust the weights in the network so that the next time that the same input is presented to the network, the network's response will be closer to the desired response. The learning procedure consists of the net starting off with a random set of weight values, choosing one of the observations or data items, and using that pattern as input, evaluating the outputs in a feed-forward manner. The errors at the output generally will be quite large, which demands changes in the weights. Using the back-propagation procedure, the net calculates $\Delta_p w_{ji}$ for all the w_{ji} in the network for that particular p. This procedure is repeated for all the patterns in the training set to yield the resulting Δw_{ji} for all the weights for that one *presentation*. The corrections to the weights are made, and the output is again evaluated in feed-forward manner. Discrepancies between actual and target output values again result in evaluation of weight changes. After complete presentation of all patterns in the training set, a new set of weights is obtained and new outputs are again evaluated in a feed-forward manner. In a successful learning exercise, the system error will decrease with the number of iterations, and the procedure will converge to a stable set of weights, which will exhibit only small fluctuations in value as further learning is attempted.

24.5. An example using Iberian Bronze Age data

I have chosen an example from Iberian Late Bronze Age to show how a back-propagation network operates. The series of warrior decorated *stelae* can be dated between 1100 BC and 7th century BC. (Barcelo 1989, Powell 1976). In this case I will be concerned with the analysis of Phoenician colonisation effects on indigenous populations, as shown by the items engraved in the *stelae*.

The network has a 3×4 matrix-configured output (Fig. 24.3), that is, three experimental variables (degree of colonisation, region and chronology), each one with 4 values:

- four chronological phases (standard chronology: Late Bronze Age II, Late Bronze Age III, Orientalising [Tartessos], Post-Colonisation [Palaeo-Iberian]),
- four geographical regions: Tajo valley, Guadiana Valley, Zujar valley, and Guadalquivir valley
- degree of colonisation, as deduced from the quantity of imported items in the Stelae: >3, >5, >7, >9.

The objective is to build a working definition of the concept COLONISATION, using as raw information the similarity among all *stelae* in the same region, and with the same chronology. In other words, I want to know if all *stelae* are similar according to an explanatory concept, and not if they are similar in iconographic terms.

This is a typical supervised-learning model. For each *stelae* we know its chronology, geographical situation and the number of engraved imported items. Input units use a 5×5 matrix to represent empirical information:

- total number of iconographic motives in a *stela*
- iconographic importance of human figure
- iconographic importance of shield
- number of "prestige" items in a *stela*
- degree of schematism

I have trained the network (in back-propagation mode) using 38 examples. In more archaeological terms, the algorithm has to calculate the relationship between iconography, geography, chronology and colonisation using those examples. The specific question I want to answer is:

> "Are the iconographically most complex *stelae* situated in the regions and chronological phases where Phoenician colonisation was stronger?"

After three hours of Computer Time (Fig. 24.4), the back-propagation engine was not able to converge to a satisfactory (in mathematical terms) solution. The best average error obtained was 0.1176, and the algorithm arrived *only* at a 96% of right answers.

How can we use these results? First of all, it is important to realise that a neural network is not an easy formula that we can use everywhere. The weight matrix can only be used by the computer. In other words, the representation of the concept COLONISATION exists only in the computer. To understand it we have not to read computer print-outs, but to begin a series of experiments with the computer solution. For instance; we may begin by comparing the input with the calculated output. Although I have stopped processing before the algorithm arrives at the tolerance level, the fit between known and calculated output is excellent. Specially good has been the correction in chronology vectors (Fig. 24.5). Known values were most of the time rather ambiguous; for example, Ecija I was dated in the Orientalising period with a 0.5 probability of error, and the probability for a post-orientalising date was also 0.5 (there were not any good chronological items). The program has re-evaluated these probabilities into 0.70 for the orientalising period and 0.30 for the post-orientalising one. The program has used the similarity between this *stelae* and others to modify the first evaluation. The same can be said for El Carneril (0.62 LBA II, 0.39 LBA III). It is also interesting the case of the Monte Blanco *stela*, a very doubtful exemplar, whose chronology has been established by the computer in the orientalising period, but with low probability rates (0.49).

After this examination we have to conclude that the network has been well trained, and that the concept (COLONISATION) has been well induced. Now we can use the concept to explain new data. I have selected a new series of 12 *stelae* (different than those in the input) as candidate instances for the concept. Validation shows many surprises. The program has produced some important mistakes trying to predict the region, number of imported items and chronology in the new set of data. Usually it confounds contiguous regions, although most of the *stelae* from the Tagus region have been correctly assigned. Chronological predictions are interesting, because the program has been able to learn only some phases (LBA II and Post-Orientalising), the other ones are not wrongly predicted, but receive a lesser probability (i.e. some items have not any sure chronology, because neither phase obtains a sufficient activation).

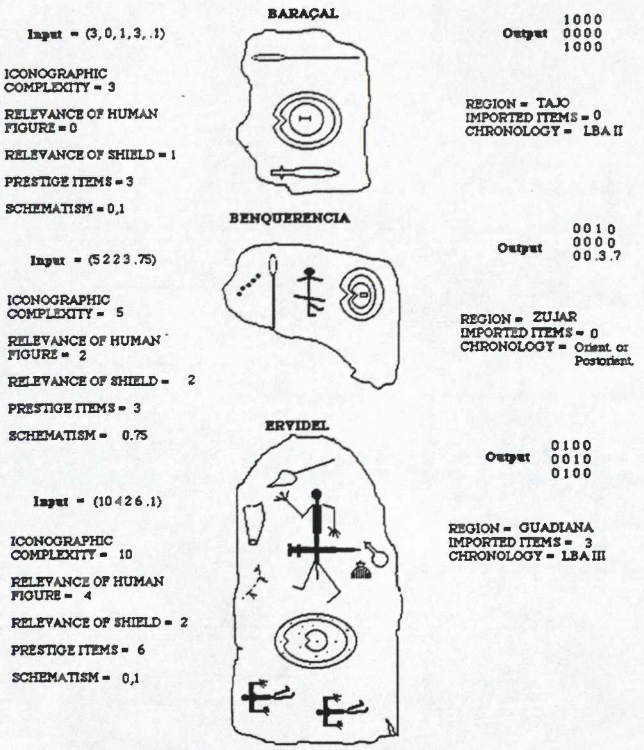


Figure 24.3: Input and output vectors for some stelae.

All these *negative* results coincide with what we expected from archaeological research in the area. Chronology can be partially explained using similarities, however the concept COLONISATION cannot be represented exclusively in terms of similarity relationships among *stelae*, because not all *stelae* in the same region, with the same chronology, and with the same number of imported items are similar in iconographic structure. In the same phase and in the same region, colonisation has different effects on the iconography of *stelae.*

24.6. Using neural network models to reproduce archaeological reasoning

A neural network or *connectionist representation* can be characterised by three general computational features: distinct layers of interconnected units, recursive rules for updating the strengths of the connections during learning, and "simple" homogeneous computing elements (Hanson and Burr 1990). Using just these three features one can construct surprisingly elegant and powerful models of conceptualisation, and even archaeological theories. What interest us here is the correspondence or mapping between numerical activity vectors and the units in the network or elements of the problem domain addressed by the model. Those elements that are subject to representation mappings can be broadly divided into two types: observable elements (input), and theoretical entities that are hypothesized but not directly observable (output) (Smolensky 1990).

The *learning* procedure is based on the strengthening of the connections among those units that co-occur and weakening the connections between pairs of units in which one is on and the other is off, and the result, a vector of activations (or matrix in the usual multidimensional case) is a *prototype.* Therefore, prototypes are represented in terms of a *response* function: the fact that a given unit is active in the output layer gives only a rough indication of the value being represented, but knowing which collection of units is active can give a quite accurate indication of the value being represented. This representation model reproduces exactly what we have analysed before. Back-propagation algorithm is supposed to reproduce the inductive mechanism, while the structure of the network (the number of units, the number of layers, the output function, the activation rule, etc.) is supposed to simulate *meta-knowledge.*

There is not any difficulty in understanding what has been represented in the input and output layers. But what is about the *hidden* layer? I have said that this layer stores the weights and interconnections between input and output; therefore, it represents the mapping between observation and theory. Consequently, if output and input units represent the definition core of any concept, then this particular mapping or correspondence represents *prototyping* relationships. That is, the degree to which we know that an item is an instance of that concept.

What is unique to the framework here described is the formal *medium* that supports representations: numerical (activity) vectors. A particular activity pattern of the network can only be defined in terms of the problem domain (input and outputs). Remember that the activation of any *hidden* layer is indicated by the activity of a set of units whose activity determines the representation of other elements in the problem domain.

Nevertheless, the prototype we can obtain using this method is not represented in linguistic terms. There is not any single list of variables and attributes. Instead we have a matrix of weights and coefficients, and it is not easy to see how we can use it for archaeological purposes. What we have obtained are "approximations" to the target units; that is, approximations to what we knew before starting the conceptualisation. As target values are the result of some controlled observation or experiment, and we have considered them as a qualitative measure of conceptualisation (1 = this input case is an instance of the given concept, 0 = this input case is not an instance of the given concept), we may use the network outputs as a *quantification* of the relationship. In other words, the network outputs are *typicality measures.* Nevertheless, we do not know the mathematical properties of these measures, nor their associated probability distribution. Shepard (1987) has pointed out that with highly similar stimuli (empirical data) or with delayed test stimuli, the relationship between similarity and distance was of Gaussian form and that the distance metric appeared to be Euclidean. This fact has not yet been analysed in neural network (although some initial work in this direction has been published by Staddon and Reid 1990).

There are some other drawbacks, too, but they are not relevant for the subject of this paper. The heart of my proposal has been to show that archaeological conceptualisation does not depend on the perceived similarities among instances, but on the results of controlled observation. Consequently we are computing a similarity relationship based

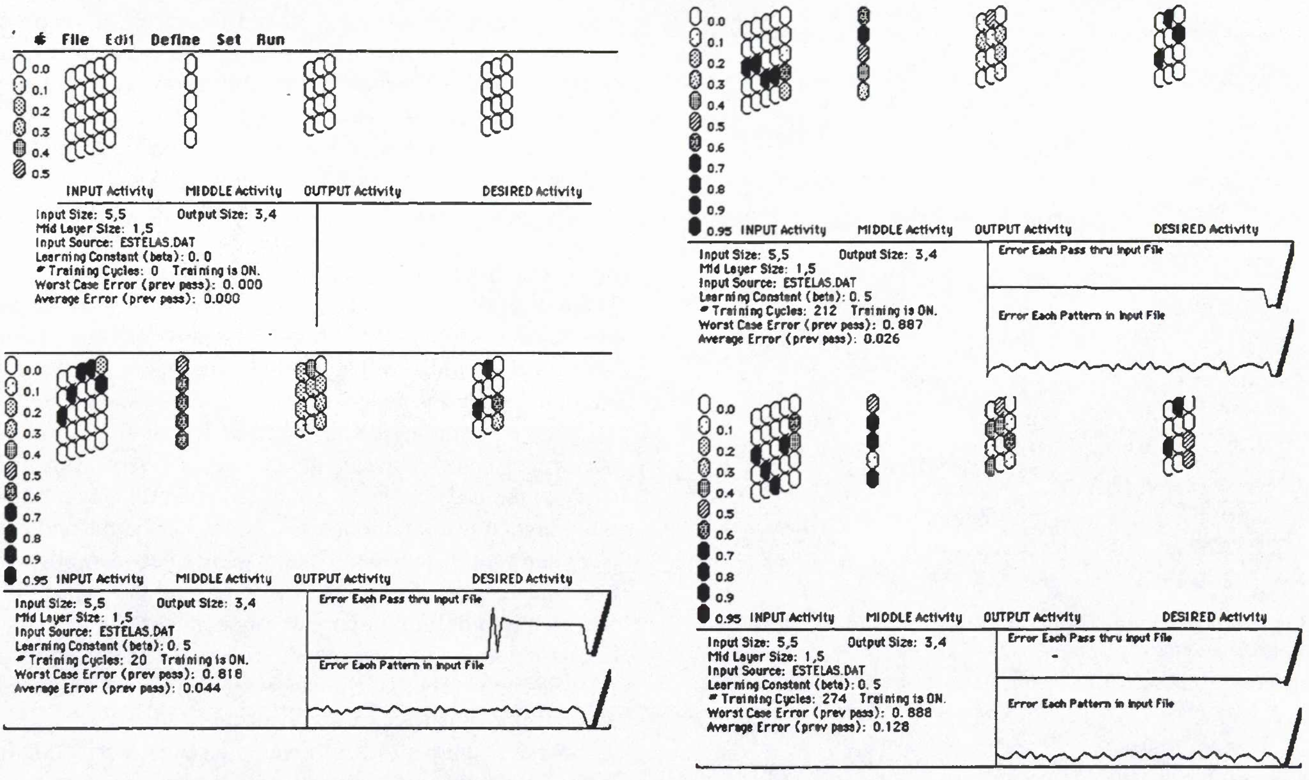


Figure 24.4: State of the network at successive training cycles.

on the identity between typicality measures: two artefacts are *similar* if and only if they are instances of the same concept. Note that *similarity* is not a consequence of "formal resemblance in general details", but of "the simultaneous activation of input units representing empirical data". Therefore, *similarity* is a consequence of typicality (the degree to which input units are instances of the same concept). Typicality is the result of the adjusting the strength of the connection between input units in proportion to the product of their simultaneous activation (*Hebbian rule*). Obviously, similar patterns of activation produce similar effects; therefore, the similarity based on "formal resemblance" appears as a special case of the model.

24.7. Conclusions

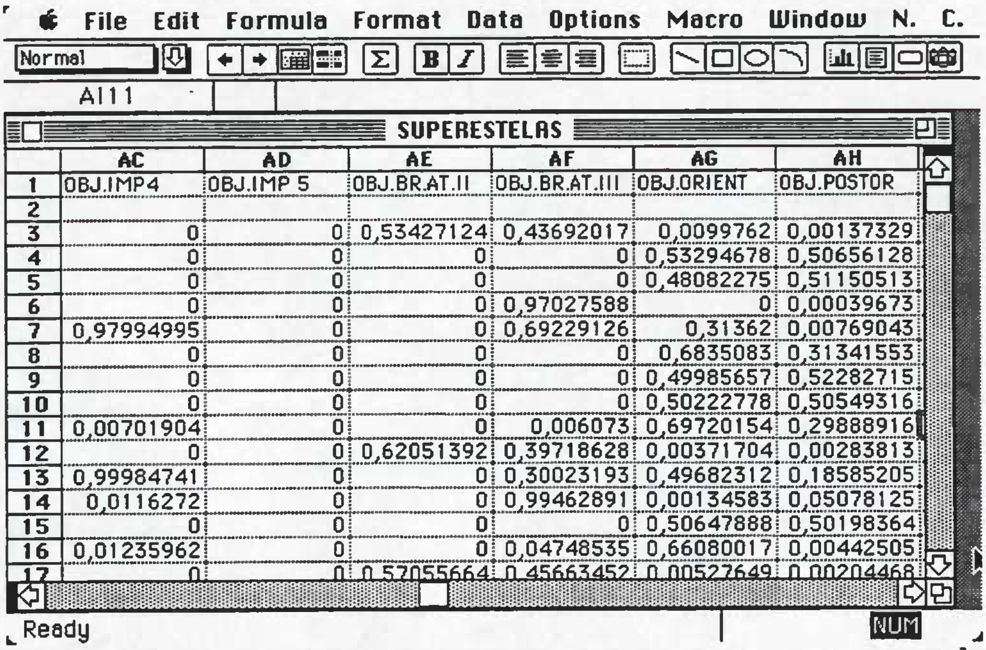
In this paper it has been show how archaeological concepts may be represented as particular patterns of activation associated with some specific inputs. In so doing, the representation of the concept is set up in such a way that the knowledge necessary influences the course of processing, and has not any meaning outside the system in which they have been produced. Consequently, it would be a serious mistake to pretend a *decoding* of the activity vector in terms of some external linguistic reality.

Archaeological concepts are theoretical entities, and not natural categories. Therefore, we have to study them by analysing how archaeologists use them, and not as intrinsic properties in the archaeological record. The main consequences we can deduce from the archaeological use of a concept are the following:

- a single concept has different instances, and there is no similarity between them.
- relationships among instances do not depend on necessary properties, but on the previous knowledge or experience we have about that concept. Differences among the instances of a concept are due to unnecessary (and non intrinsic) properties.
- instead of offering defining conditions, concepts are intrinsically *fuzzy*. Their definition core depends on the way we use the concepts.

Neural network systems and back-propagation algorithms allow archaeologists to follow these definition assumptions. Any stable pattern of activity is a concept, but a stable pattern of activity is only that whose units satisfy some of the micro-inferences and violate others. That is to say, a stable pattern of activity is one that violates the plausible micro-inferences less than any of the neighbouring patterns. We may modify the concept representation schema by changing the inference rules so that the new pattern violates them less than its neighbours. We have to consider archaeological conceptualisation as a process that constructs a pattern of activity, which represents the most plausible item that is consistent with the given knowledge. Note that we are speaking about the *most* plausible, and not about a structurally correct one. Neural networks provide us with good but not optimal fit.

Some authors (see, Partridge 1990, among others) have stressed the inappropriateness of neural networks to simulate scientific reasoning. The main drawback of this model is its *black box* character, preventing a description of the system in linguistic terms. I suppose that most archaeologists may see it as a fundamental drawback, on the opposite I think that this "non-linguistic" character of representation is one of the main advantages of the model (see also Gardin 1990). Because of the supervised-learn-



	AC	AD	AE	AF	AG	AH
1	OBJ.IMP4	OBJ.IMP 5	OBJ.BR.AT.II	OBJ.BR.AT.III	OBJ.ORIENT	OBJ.POSTOR
2						
3	0	0	0,53427124	0,43692017	0,0099762	0,00137329
4	0	0	0	0	0,53294678	0,50656128
5	0	0	0	0	0,48082275	0,51150513
6	0	0	0	0,97027588	0	0,00039673
7	0,97994995	0	0	0,69229126	0,31362	0,00769043
8	0	0	0	0	0,6835083	0,31341553
9	0	0	0	0	0,49985657	0,52282715
10	0	0	0	0	0,50222778	0,50549316
11	0,00701904	0	0	0,006073	0,69720154	0,29888916
12	0	0	0,62051392	0,39718628	0,00371704	0,00283813
13	0,99984741	0	0	0,30023193	0,49682312	0,18585205
14	0,0116272	0	0	0,99462891	0,00134583	0,05078125
15	0	0	0	0	0,50647888	0,50198364
16	0,01235962	0	0	0,04748535	0,66080017	0,00442505
17	0	0	0,57055664	0,45663452	0,00527649	0,00204468

Figure 24.5: Chronology as predicted by the neural network (expressed in probability terms).

ing philosophy, concepts represented in a neural network have a meaning according to examples used in its construction. Probably it cannot be easily expressed in linguistic terms, but they are useful using a powerful computer. Nowadays, it is common to say that the theoretical literature on similarity and classification has diverged from archaeological practice to such a degree that the two are now unrelated. In this paper I have tried to solve this disparity. It may seem contradictory that a *black box* can be a practical system. The question is how to use a neural network, and not why to select complex procedures to solve simple problems. Archaeological problems are not simple, although some archaeologists seem to think so.

The future evolution of statistical research in archaeology seems to be in this new domain. The goal of this paper is only to show some new methods, which can be used to integrate theory building and statistical discovery.

Bibliography

ADAMS, W. Y. 1988. "Archaeological classification: theory versus practice", *Antiquity* 61: 40–56.

ANDERSON, J. R. 1990. *The adaptative character of thought.* Erlbaum, Hillsdale, New Jersey.

ASHBY, F. G. & W. W. LEE, 1991. "Predicting similarity and categorization from identification", *Journal of Experimental Psychology: General.* 120(2): 150–172.

ASHBY, F. G. & N. A. PERRIN, 1988. "Toward a Unified Theory of Similarity and Recognition", *Psychological Review* 95: 124–150.

BARCELO, J. A. 1989. *Arqueologia, Logica y Estadistica. Un Analisis de las Estelas de la Edad del Bronce en la Peninsula Iberica,* Universidad Autonoma de Barcelona, Ph.D. Thesis.

BARCELO, J. A. 1993a. "Automatic Problem Solving in Archaeology: a computational framework", *Archeologia e Calcolatori* 4 (in press).

BARCELO, J. A. 1993b. *Arqueologia Automáaica. La utilizacion de tecnicas de Inteligencia Artificial en Arqueologia Teorica,* Editorial Critica, Barcelona.

BARSALOU, L. W. 1989. "Intraconcept similarity and its implications for interconcept similarity", *in* Vossniadou, S. & A. Ortony (eds), *Similarity and Analogical Reasoning.* Cambridge University Press, Cambridge.

CHURCHLAND, P. M. 1988. *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science.* The MIT Press, Cambridge, Massachussets.

COWGILL, G. L. 1990. "Artifact classification and archaeological purposes", *in* Voorrips, A. (ed.), *Mathematics and Information Science in Archaeology: A Flexible Framework.* Studies in Modern Archaeology 3, Holos-Verlag, Bonn.

ENNIS, D. M., J. J. PALEN & K. MULLEN 1988. "A multidimensional stochastic theory of similarity", *Journal of Mathematical Psychology* 32: 449–465.

FISHER, D. & D. PAZZANI 1991. "Computational Models of Concept Learning", *in* Fisher, D. H., M. J. Pazzani & P. Langley (eds), *Concept Formation: Knowledge and Experience in Unsupervised Learning.* Morgan Kaufmann, San Mateo, California.

FORSYTH, R. (ed.) 1989. *Machine Learning. Principles and Techniques.* Chapman and Hall, London.

GARDIN, J. C. 1990. "The structure of archaeological theories", *in* Voorrips, A. (ed.), *Mathematics and Information Science: a flexible framework.* Studies in Modern Archaeology 3, Holos-Verlag, Bonn.

GLUCKSBERG, S. & B. KEYSAR 1990. "Understanding metaphorical comparisons: Beyond Similarity", *Psychological Review* 97: 3–18.

GOODING, D. 1990. *Experiment and the making of Meaning.* Kluwer, Dordrecht, The Netherlands.

HANSON, S. J. & D. J. BURR 1990. "What connectionist models learn: Learning and representation in connectionist networks", *Behavioral and Brain Sciences* 13: 471–518.

HODSON, F. 1980. "Cultures as types? Some elements of classification theory", *University of London Institute of Archaeology Bulletin* 17.

HOOKER, C. A. 1987. *A realistic theory of science.* State University of New York Press, Albany.

KURMHANSL, C. L. 1978. "Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density", *Psychological Review* 85: 445–463.

MICHALSKI, R. S. 1989. "Two-tiered concept meaning, inferential matching and conceptual cohesiveness", *in* Vossniadou, S. & A. Ortony (eds), *Similarity and Analogical Reasoning*. Cambridge University Press, Cambridge.

MURPHY, G. L. & D. L. MEDIN 1985. "The role of theories in conceptual coherence", *Psychological Review* 92: 289–316.

NIINILUOTO, I. & R. TUOMELA (eds) 1973. *Theoretical concepts and hypothetico-inductive inference*. Reidel, Dordrecht, The Netherlands.

NOSOFSKY, R. M. 1988. "Similarity, frequency and category representations", *Journal of Experimental Psychology: Learning, Memory and Cognition* 14: 54–65.

NOSOFSKY, R. M. 1992. "Similarity scaling and Cognitive Process Models", *Annual Review of Psychology* 43: 25–53.

ORTONY, A. 1979. "Beyond literal similarity", *Psychological Review* 86: 161–180.

OSHERSON, D. N. 1987. "New Axioms for the Contrast Model of Similarity", *Journal of Mathematical Psychology*, 31: 93–103.

PAO, Y. H. 1989. *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, Reading, Massachussets.

PARTRIDGE, D. 1990,."Connectionism is better for Engineers than for Scientists", *in* Tiles, J. E., G. T. MacKee, & G. C. Dean (eds), *Evolving Knowledge in Natural Science and Artificial Intelligence*. Pitman, London.

PITRAT, J. 1990. *Métaconnaissance. Futur de l' Intelligence Artificielle*. Hermès, Paris.

POWELL, T. G. E. 1976. "South-western Iberian peninsular chariot stelae". *in* Megaw, J. V. S. (ed.), *To illustrate the monuments. Essays on archaeology presented to Stuart Piggott*. Thames and Hudson, London.

READ, D. W. 1989. "Intuitive typology and Automatic Classification: Divergence or Full Circle?", *Journal of Anthropological Archaeology* 8: 155–188.

RIPS, L. 1989. "Similarity, typicality, and categorization", *in* Vossniadou, S. & A. Ortony (eds), *Similarity and Analogical Reasoning*. Cambridge University Press, Cambridge.

RIVADULLA, A. 1986. *Filosofia actual de la Ciencia*. Tecnos, Madrid.

ROSCH, E. & C. B. MERVIS 1975. "Family resemblances. Studies in the internal structure of categories", *Cognitive Psychology* 7: 573–605.

RUMELHART, D. E., G. HINTON & R. J. WILLIAMS 1986. "Learning Internal Representation by Error Propagation", *in* Rumelhart, D. E., J. L. McClelland & the PDP Group (eds), *Parallel Distributed Processing* 1, The MIT Press, Cambridge, Massachussets.

SHEPARD, R. N. 1987. "Toward a universal law of generalization for psychological science", *Science* 237: 1317–1323.

SHOEMAKER, S. S. 1980. "Properties, Causation and Projectability", *in* Cohen, L. J. & M. Hesse (eds), *Applications of Inductive Logic*. Clarendon Press, Oxford.

SMITH, L. B. 1989. "From global similarities to kinds of similarities: the construction of dimensions in development", *in* Vossniadou, S. & A. Ortony (eds), *Similarity and Analogical Reasoning*. Cambridge University Press, Cambridge.

SMOLENSKY, P. 1990. "Representation in connectionist networks", *Intellectica* 9–10: 127–165.

SOKAL, R. R. 1988. "Unsolved problems in numerical taxonomy", *in* Bock, H. H. (ed.), *Classification and related methods of data analysis*. North Holland, Amsterdam.

STADDON, J. E. R. & A. K. REID 1990. "On the dynamics of generalization", *Psychological Review* 97: 576–578.

STEPP, R. E. & R. S. MICHALSKI 1986. "Conceptual Clustering of Structured Objects: a Goal-Oriented Approach", *Artificial Intelligence* 28: 43–69.

STERNBERG, R. J. & E. E. SMITH (eds) 1988. *The Psychology of Human Thought*. Cambridge University Press, Cambridge.

STONE, G. O. 1986. "An analysis of the Delta Rule and the Learning of Statistical Associations", *in* Rumelhart, D. E., J. L. McClelland & the PDP Group (eds), *Parallel Distributed Processing* 1. The MIT Press, Cambridge, Massachussets.

THAGARD, P. 1988. *Computational Philosophy of Science*. The MIT Press, Cambridge, Massachussets.

THAGARD, P. 1989. "Explanatory Coherence", *Behavioral and Brain Sciences* 12: 435–502.

TUOMELA, R. 1973. *Theoretical Concepts* 10. Springer-Verlag, Vienna and New York.

TVERSKY, B. 1977. "Features on Similarity", *Psychological Review* 85: 327–352.

WHALLON, R. & J. A. BROWN (eds) 1982. *Essays on archaeological typology*. Center for American Archaeology Press, Evanston Illinois.

WATTENMAKER, W. D., G. V. NAKAMURA & D. L. MEDIN 1988. "Relationships between similarity-based and Explanation-based categorization", *in* Hilton, D. J. (ed.), *Contemporary Science and Natural Explanation. Commonsense conceptions of causality*, The Harvester Press, Brighton.

ZEIDENBERG, M. 1990. *Neural Networks in Artificial Intelligence*. Ellis Horwood, Chichester.

Dr. Juan A. Barcelo
Universitat Autonoma de Barcelona
Dpt. Historia de les Societats Precapitalistes i Antropologia Social
Facultat de Lletres. Edifici B.
E-08193 Bellaterra