Archeologia e Calcolatori
7, 1996, 561-577

UNDERPINNING THE DISCIPLINE ONE HUNDRED YEARS (OR MORE) OF CLASSIFICATION IN ARCHAEOLOGY

1. Introduction

In 1995, we celebrate the publication of one of the most seminal papers in archaeology: Dragendorff's work on samian pottery (Dragendorff 1895). It has become one of the most-quoted archaeological papers of all time, although the full reference rarely appears in bibliographies. Indeed, one might say that the mark of true success is for a paper to be quoted so widely that there is no need to give the reference. Why was the paper so successful? There are two reasons:

(i) it presented a style of drawing which rapidly became a world-wide standard, with only minor regional variations, and
(ii) it established the idea of a pottery type series as a useful tool for the presentation of pottery from archaeological excavations or collections.

It would be fascinating to trace the history of the development of drawing styles, and their relationship to archaeological concerns and theory, but here we must concentrate on the second point, the idea of types, type series, and the classification of artefacts.

Dragendorff was not the first to construct a pottery type series; he was preceded by Plique (1887), also working on samian (from Lezoux), and, it has been claimed by Rhodes (1979, 89), by Roach Smith working in London (Smith 1854; 1859). But in innovation it is not necessarily the first occurrence that counts, but the first one that is widely adopted by its community. It is on these grounds that I make a claim on behalf of Dragendorff's work, which is still in everyday use today. Why was his type series so successful? One reason is that it was so practically useful: it dealt with a class of material (samian pottery) that is found in many places in embarrassingly large quantities, which was (for archaeological material) highly standardised and therefore lent itself to a classificatory approach. The type series made it simple to sort further pots, or even quite small sherds, into their respective types, and it was capable of extension: later workers in the field (e.g. Déchelette, Ludowici, Knorr, Walters) added their own supplementary type series.

One must not give the impression that classification and typology were, even at first, only about pottery. Pitt-Rivers had lectured on *The Principles of Classification* as early as 1874 (published in Pitt-Rivers 1906). Flinders Petrie (1904) was also working on the theoretical aspects of classification, and this work can be paralleled elsewhere in Europe, e.g. by Montelius (1903). It was very much in the 19th-century way of thought that classification was a way of imposing order on chaos, thus gaining power over it and hopefully learning something about it (though the last was sometimes lost sight of).

Here, I shall first look at classification from an archaeologist's point of view – its uses and the practical issues – then from a mathematician's point of view – what it is and how it works. I shall then look at historical trends from DRAGENDORFF to the present day, before trying to pick out points for the future. If I show a bias towards pottery studies, as against flints or other classes of artefact, that is because pottery is the material of which I have the most experience.

2. What is classification for?

Classification as an activity is deeply embedded in the archaeological psyche. For example, CHANG (1967, 71) claimed that it occupied 80 to 90% of archaeologists' time, and archaeologists have for many years debated the purposes and practices of classification (the literature is vast; see, for example, the bibliography in BAXTER 1994). Even if Chang's figure is an exaggeration, one hopes that this time is being spent to good purpose, and is not just a vast collective habit. So what is classification for, in a practical sense? Four reasons can be suggested:

(i) economy of reporting,
(ii) economy of thought,
(iii) generalisation of evidence,
(iv) generation and testing of hypotheses.

2.1 *Economy of reporting*

This is a very practical reason. As archaeologists, we must report our findings to our colleagues, and to the general public. One of the many obstacles to publication is the sheer volume of artefacts that an excavation may generate – thousands of pots, or of flints, or hundreds of objects of other classes, such as metalwork. If we had to draw and/or describe every sherd, or every flint, individually, we would never finish in our own lifetimes (some never do, but that is another subject). One answer to the problem is to recognise, as Dragendorff did, that many of our objects are remarkably similar, and some are almost identical. That being so, it is a waste of time to draw or describe them all: one can present a representative example (the 'type specimen') and refer the others of its type to it (or simply say "so many examples of type such-and-such").

A good example of the benefits of this approach comes from the study of pottery fabrics. Until the 1970s, British pottery reports tended to contain many very brief, and virtually useless, descriptions, such as "buff sandy ware". Now they are likely to contain a single long and detailed description of each ware present (or a reference to a description published elsewhere) and quantified catalogues of sherds of each ware and form. Modern printing permits the publication of high-quality photographs (e.g. Davies *et al.* 1994, Plate 5), which may even make the written descriptions redundant.

2.2 *Economy of thought*

Another fundamental archaeological activity is that of comparison; we compare one pot with another, one assemblage with another, one site with another, and so on. We ask questions like "is the pottery found at site A the same as that at site B?" or, more usefully, "are the assemblages of pottery at site A like the assemblages at site B?" Answers to such questions enable us to address issues of chronology, distribution and trade, or status and/or function.

Comparing even two small assemblages on a sherd-by-sherd basis would be a tedious and probably impossible task. We would have to physically bring them together, or at least try to remember what one was like while examining the other, since the terse "buff sandy ware"-type descriptions are totally inadequate for comparative purposes. But if one assemblage had been divided into types, we could use the detailed descriptions (or, better still, actual examples of types) to see whether the same types occur in the other assemblage. This gives us a basis for comparison while preserving our sanity, since we do not have to remember vast amounts of individual detail.

2.3 *Generalisation of evidence*

Different contexts yield different amounts, and different sorts, of information about the artefacts found in them. For example, some contexts are securely dated and thus provide dating evidence, while others do not; some contexts provide rich assemblages of artefact types which may tell us about function or status of those types, while others contain very few types. Production sites may give evidence for the manufacture of certain types, which knowledge can be transferred to consumption sites to give evidence for trade or other means of distribution.

The point about all these examples is that, because we can recognise the same types in different contexts, we can transfer information about one, to the other. This is the basis, for example, of cross-dating, which has been used since Montelius and Petrie in the late 19th century (Graslund 1987, Drower 1985 respectively). The dangers of uncritical use of this approach are well known: an artefact found in a dated context may be at an extreme of its type's date-range, or even residual (i.e. outside its type's usage date-range), and the transfer of that context's date to another may therefore be misleading. Another danger is that types (e.g. form types) may diffuse outwards from a centre as a wave of innovation (like ripples on a lake), so that a type in a peripheral region may be neither manufactured nor in use contemporaneously with the same type at the centre. The transfer of other types of evidence (e.g. source, status, function) can pose similar dangers.

An important point that emerges from this discussion is that we cannot expect there to be a single 'all-purpose' typology for a class of artefacts. We need different sorts of typologies for different purposes. For example, if our main use of the class was for dating purposes, we would need a chronological typology, which might be based on stylistic aspects; if we were interested mainly in distribution we would need a typology that reflected source (e.g. chemical or physical composition); while if function was our interest, we would need a functional typology (perhaps based on less obvious characteristics, such as use-wear, or the interpretation of complete assemblages).

2.4 Generation and testing of hypothesis

Once we are able to observe chronological or spatial patterns of types (whether on a small or large scale), and of assemblages in terms of their component types, we can start to ask, and hope to answer, more interesting and deeper questions. A secure chronology, for example, enables us to ask questions about the direction and rate of social or technological change. It may be argued that modern scientific dating techniques have removed the need for artefact-based dating. While there is some truth in this, especially in the broader picture, there are many questions for which (e.g.) radiocarbon dating, with standard deviations of perhaps 50 to 100 years, simply do not provide fine enough chronological resolution (Bayliss, Orton 1994).

Spatial patterns at a regional level enable us to address questions of the mode of distribution. Is the presence of artefacts away from their centre of production due to 'trade', 'redistributive exchange', or even (not so common as an explanation today) 'invasion'? At a smaller scale, patterning within sites may shed light on questions of zoning, functional or 'activity' areas, or the relationship between (for example) different social classes within an urban settlement.

The basic point that I have tried to make here is that all these interesting archaeological activities are only possible once we have a secure typological foundation (or foundations) for our work. If we have types, we have comparisons and patterns, which can lead to interpretation; if all objects are unique, there is no operational basis for comparison and hence the observation of pattern, and then interpretation, is not possible.

3. What is classification?

Having seen why archaeologists need to classify artefacts, we now look at how this activity appears to those who study classification in a more formal and abstract sense. To an archaeologist, classification may be an intuitive exercise, almost as natural as breathing, but for that very reason not easily studied or transferred. One danger is that a specialist's expertise may die with them if it cannot be expressed to another archaeologist; a lesser but very real problem is that of bottlenecks in some specialisations (in Britain an excavation report may be delayed for years awaiting a report from a samian or mortarium specialist). A formal study of classification may help to elucidate the mental processes involved, or even replicate or supplement them on a computer, perhaps by means of an expert system, thus enabling skills and expertise to be more readily transferred.

From a mathematical point of view, objects (including artefacts) may be seen as points in a mathematical space (the 'object-space' or 'description-space'). The space may be defined explicitly by the variables which form its axes or dimensions, in which case each point is defined by its values on those variables. For example, if we consider that length, breadth and depth are the only variables worth recording, then we have a three-dimensional space with each object represented by a point located at the values of its length, breadth and depth. An alternative approach is to define the space by the distances between the pairs of points in it. If we choose the former approach, we have the simple and familiar 'data matrix' in which each column represents a variable and each row an object. In the latter case, we have a matrix (usually, but not necessarily, symmetric about its diagonal) of inter-point distances or similarities. An example of the latter approach would be to attempt to recreate a map from the distances between each pair of a set of towns.

The process of classification then consists of dividing up our space ('partitioning' it) into subsets which represent types, in a way that is both mathematically valid and practically useful. Five criteria are commonly used to assess the value or usefulness of such a division:

<u>mathematical</u>

(i) every object should belong to a type (the 'exhaustive' criterion),
(ii) no object should belong to more than one type (the 'exclusive' criterion),
(iii) given a new object, it should be possible to assign it to a type,

<u>archaeological</u>

(iv) all the objects in a type should be more like each other than they are like those in other types (the types should be 'natural'),
(v) the assignment in (iii) should be easy to do.

The first three are formal criteria which describe the process. While it is possible to devise schemes which do not follow them, trouble or confusion usually results. The last two are more practical, and their realisation depends on the data as well as the process. For example, if there are no 'natural' types, no process will find them.

There are two main broad approaches to the process of classification, the explicit and the implicit.

(a) explicit: we seek to divide the space by creating boundaries ('frontiers') in terms of formal definitions. For example, we might define the pottery form types 'jar', 'bowl' 'dish' and 'plate' in terms of the height/diameter ratio (jars have a ratio > 1, etc., see Fig. 1).

This satisfies all the criteria except possibly (iv), since it is quite possible to define a frontier through a natural group rather than between them (see Fig. 2). It may seem odd archaeologically, in that types are defined in terms of their 'edges', where there are no or few examples, and those that are there are not representative of the type as a whole.

(b) implicit: types are defined in terms of their members, or of a 'typical' member and others thought to be similar to it. This often results in more 'natural' groupings than approach (a), but it can be difficult to fit new object into an existing typology. In extreme circumstances, one new object may alter the whole structure (HODSON 1970, 307), although it has been suggested that this problem may be more theoretical than real (WRIGHT 1989).

3.1 *Implementation*

A vast range of techniques is available for implementing these approaches; even in the 1970s the literature on the subject was vast (CORMACK 1971). There is the whole family of cluster analyses (e.g. EVERITT 1980; GORDON 1987) – single-link, average-link, Ward's method, k-means, etc. – which would require a separate paper to describe in detail individually. They have recently been assessed in an archaeological context by BAXTER (1994, 140-184). They suffer from the common failing that they create groupings ('clusters') whether or not such exist naturally in the data – it is the analyses' job to do so. For that reason, it may be desirable to supplement them with a technique that simply displays the data without attempting to group it, for example principal components analysis (pca), or correspondence analysis (ca) (*ibid.*, 48-99, 100-139 respectively). A very interesting approach to displaying data relating to complex objects is 'inexact graph matching' (DALLAS 1992; see below).

3.2 *Choice of variables*

This mathematical model of classification reinforces the point made earlier, that we cannot expect there to be a single all-purpose typology for any set of archaeological objects. In mathematical terms, a particular set of objects has an unbounded number of variables (since we could always think of another aspect to record or measure, if we wished to do so), and therefore selection of the variables to be used is inevitable. Different sets of variables will define different spaces, within which the patterns of relationships between the variables may be quite different. In other words, definition of types

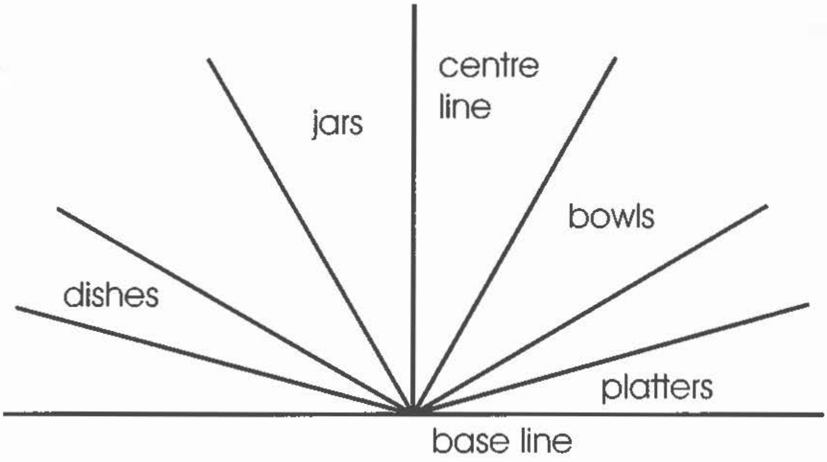


Fig. 1 – Example of template for assigning pottery vessel shapes to jar, bowl, dish or platter classes on the basis of their height/diameter ratio.

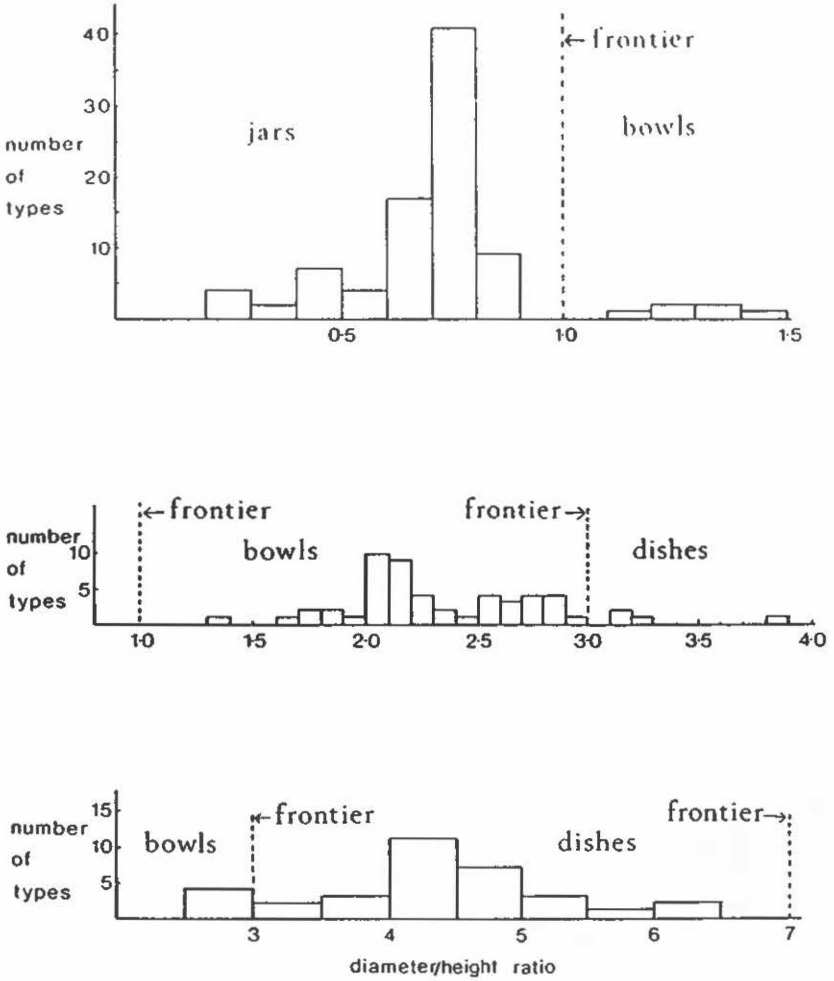


Fig. 2 – Histograms showing the height/diameter ratios of pottery vessel shapes classified by GILLAM (1970) as jars, bowls and dishes respectively.

depends crucially on the investigator's choice of variables.

If the variables are of different types (e.g. nominal, ordinal), the relative weightings between them may not be apparent, and arbitrary decisions about relative weights have to be made. These too can radically affect the geometry of the space in which the objects are located. The Gower coefficient (DORAN, HODSON 1975, 142-3), which combines variables of different types into a single similarity coefficient, is sometimes advocated as a way of avoiding such problems. While it does so in purely operational terms, it contains its own arbitrary decisions, and is notorious, for example, for tending to give more weight to discrete than to continuous variables.

A mathematical analysis of classification, therefore, does not make classification an 'objective' exercise. Rather, it exposes the subjective elements within it. The hope is that by separating the subjective input from the objective formal process of classification, we can improve both the definitions of archaeological objectives (the input) and the methodology of classification itself.

3.3 *Levels of data*

It is worth pointing out at this stage that classification can take place on many levels, e.g.

(i) individual objects,
(ii) assemblages,
(iii) chemical compositions.

Of these, (i) has traditionally been the area of the individual specialist, and in some fields this is still very much the case. By contrast, (iii) with its large tables of data, cries out for a computerised statistical approach, and much work has been done on assessing the suitability of various techniques and approaches (BAXTER 1994). It is perhaps (ii) that needs the most immediate attention: it is difficult for the individual to characterise and compare whole assemblages, and while statistical techniques are being developed (e.g. ORTON, TYERS 1992), their use is not yet widespread. One major obstacle is the difficulty of making the outcomes of such analyses accessible to the 'ordinary' archaeologist; this is an area to which effort should be directed.

4. HISTORICAL TRENDS

The 'intuitive' phase of classification in archaeology persisted from its inception until the 1960s. There were theoretical developments and debates within this phase, for example on the use of two-tier classificatory systems (KRIEGER 1944) and the use of 'modes' rather than 'attributes' as a basis for classification (ROUSE 1960).

In the 1960s, archaeology came under the influence of the 'numerical taxonomy' school of thought, *via* the work of the microbiologist Peter Sneath (SOKAL, SNEATH 1963; HODSON *et al.* 1966). The late 1960s and early 1970s were a period of intensive experimentation with a wide range of datasets and techniques (e.g. Fig. 3), the conclusions of which were well summarised by DORAN and HODSON (1975, 158-186).

Inevitably, there was a backlash as some of the experiments were seen to 'fail', and the value of quantitative approaches in this area was challenged. At about the same time, the 'typological' approach to archaeology, and to pottery studies in particular, was seriously questioned, partly in response to the perceived aridity of much of the earlier 'cultural-historical' approach (e.g. by DE VORE 1968), and partly in response to new ideas coming from workers like SHEPARD (1956). As so often happens, a more balanced 'middle' view emerged from the extremes (e.g. ALDENDERFER 1987).

Approaches other than the multivariate analysis of two-way data tables have been examined. One such was the use of expert systems (e.g. BAKER 1988). The laudable aim of this approach was to by-pass the bottleneck in archaeological research caused by a dearth of specialists in certain key areas (e.g. samian pottery), by encapsulating the specialist's knowledge in a computer program. It was further hoped that this process would shed light on the process of classification itself (as well as replicating the specialists' outcomes). This effort has been mainly disappointing, and relatively little has been achieved beyond what could be achieved by more conventional means. The problems were perhaps even more difficult than had been thought, or perhaps archaeologists have been trying to tackle the wrong questions.

More promising perhaps is the approach known as 'inexact graph-matching', used by DALLAS (1992). This can be said to bear the same relationship to conventional 'data table' statistical analysis as relational databases do to flat-file databases or to spreadsheets. It represents complex objects as "a graph structure in which the parts correspond to nodes, labelled with lists of property values, and the arcs are labelled with lists of relationship values" (*ibid.*, 171). Such a model can be used, for example, to describe the ground plan of a house (compare Figs. 4 and 5), or a sculpture, or a grave. Similarities between corresponding nodes (e.g. a particular room in a house, or a particular class of artefact in a grave) can be studied by conventional means, but in addition the similarities between graph structures can be studied, and the two combined to create overall similarities between complex objects. Although there are arbitrary decisions hidden in the way that the similarities between elements, and between graph structures, are combined, this seems to be a very promising approach.

There has been a trend within multivariate statistics towards the more visual presentation of data. It can be partly linked to the 'exploratory data analysis' (EDA) school of thought (TUKEY 1977), partly to the increasing power and sophistication of computers, and partly to the demands from archaeolo-

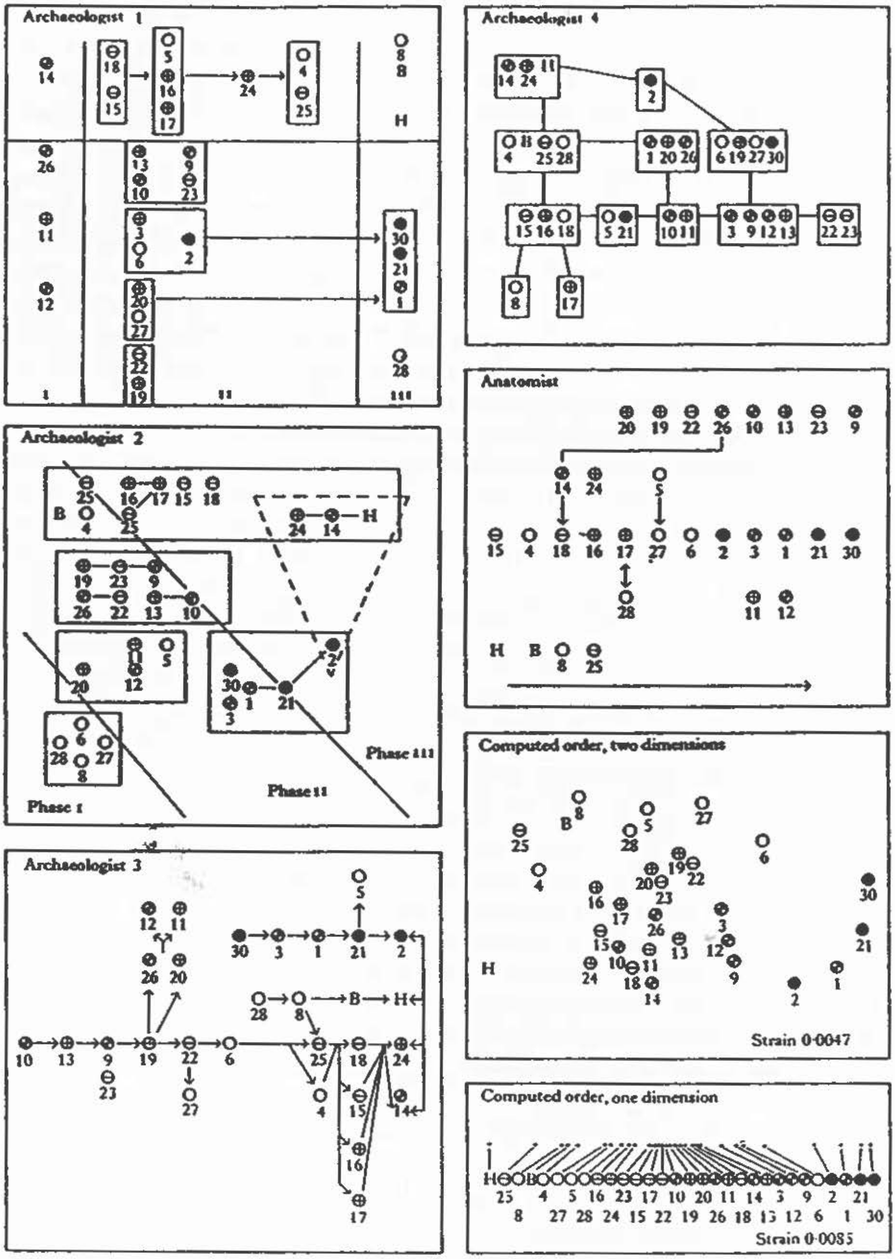
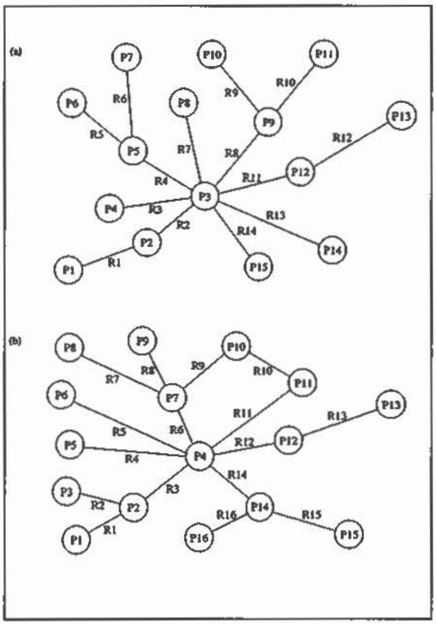


Fig. 3 – Comparison of intuitive and computed classifications of a test sample of thirty fibulae. Computed classifications are given by non-metric multidimensional scaling (HODSON *et al.* 1966, 320).



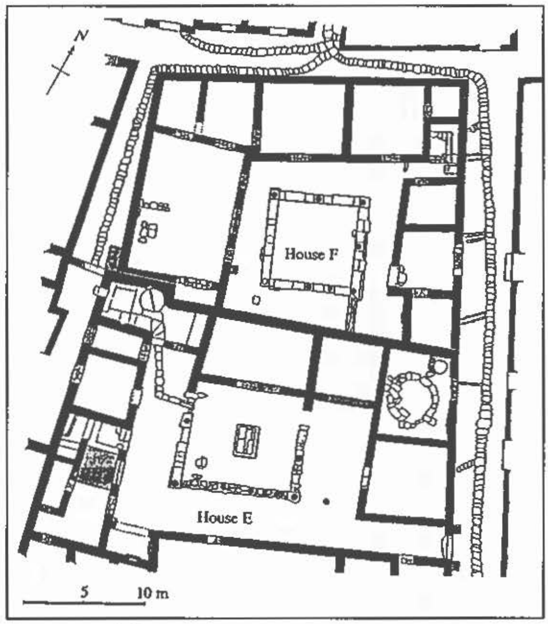


Fig. 4 – Relational description graphs for the ground floor plans of Delos houses IIE and IIF (DALLAS 1992, Fig. 22.2).

Fig. 5 – Ground floor plans of the Delos houses IIE and IIF (DALLAS 1992, Fig. 22.1).

gists and other users for more accessible presentation of their data. This demand is understandable, since most archaeologists find it easier to interpret data presented in a graphical rather than a tabular form, but has its dangers, since the human eye is very good at seeing patterns, even when they do not exist (HODDER, ORTON 1976, 4-6).

One area that has seen considerable development in the past 30 years is the study of shape (e.g. of pots, but also of other classes of artefact, such as axes). Attempts to produce an 'objective' classification of shape started with the use of simple ratios (e.g. of height to diameter) (WEBSTER 1964; HARDY-SMITH 1974). More of the profile of the vessel was used in the 'sliced' method (WILCOCK, SHENNAN 1975, see Fig. 6): the diameters at intervals up the profile of the pot, expressed as a ratio (e.g. of height) were used as input to a conventional multivariate technique, such as pca or cluster analysis. A good example of this approach is RICHARDS' (1987) study of Anglo-Saxon burial urns. A related approach is the 'swept-radius' method (LIMING *et al.* 1989). For technical reasons, it may be easier to compare and classify mathematical transformations of the shape of an artefact, rather than the shape itself. Such thinking underlies a range of techniques, e.g. the tangent-profile (TP) technique (MAIN 1981; LEESE, MAIN 1983; see Fig. 7), the sampled tangent-profile (STP) technique (MAIN 1986), B-spline curves (HALL, LAFLIN 1984), Fourier

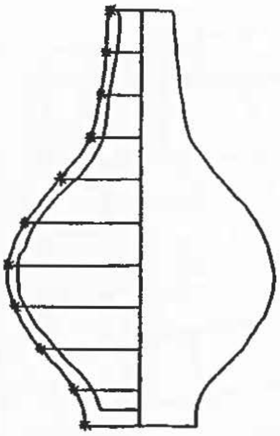


Fig. 6 – Example of the 'sliced' method of describing the shape of a pottery vessel (SHENNAN and WILCOCK 1975).

series (GERO, MAZZULLO 1984), the centroid and cyclical curve technique (TYLDESLEY *et al.* 1985) and the two-curve system (HAGSTRUM, HILDEBRAND 1990). The most recent suggestion is the Generalised Hough Transform (DURHAM *et al.* 1995), which can match parts of a shape (e.g. sherds) as well as complete shapes.

An approach that has recently been making an impact on classification, among many areas of archaeological study, is the use of bayesian statistical analysis (BUCK, LITTON 1991). This can be summed up briefly as the application of the equation

prior belief + data = posterior belief,

i.e. the "orderly influencing of opinions by data". Although to some extent controversial (this is not the right place to go into the controversy), it is becoming accepted as a useful and flexible tool, capable of integrating different sorts of data (e.g. stratigraphy and scientific dating determinations). What prior beliefs might contribute to a classificatory study? Some examples are the likely number, size or shape of clusters in a cluster analysis (*ibid.*, 94), or data from earlier work done on similar material. Heavy demands on computer resources, and 'difficult' mathematics mean that adoption of this approach is not likely to be rapid, but we can expect to see it spread as its benefits are recognised and obstacles overcome.

5. PRESENT STATE AND A LOOK TO THE FUTURE

Recent developments in statistical software (e.g. SPSS-PC) have enabled sophisticated classificatory software to be run on quite modest desktop PCs, within the range of most academics and many archaeologists. The temptation therefore is to gratefully accept the tools provided by the commercial

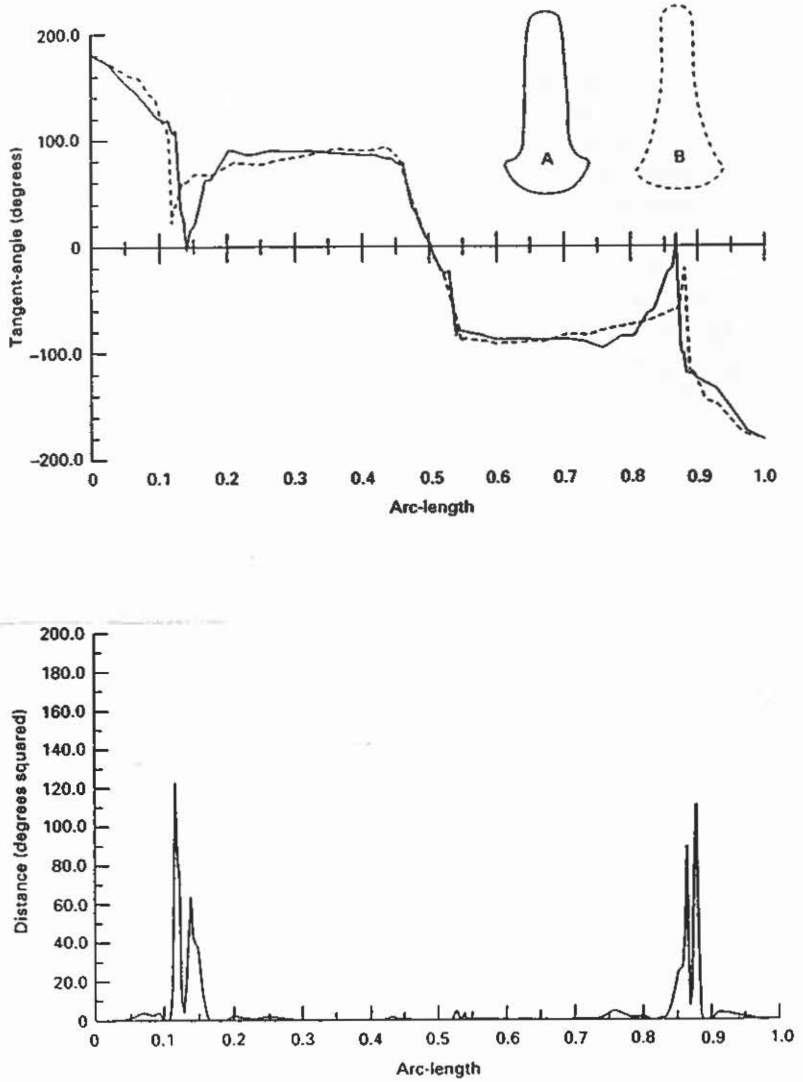


Fig. 7 – Example of tangent-profile descriptions for two axes, and the distance between them (LEESE, MAIN 1983, fig. 3).

world (e.g. the many forms of cluster analysis, principal components analysis, correspondence analysis, etc.), and not to worry too much about what goes on 'inside the box'. This would be a mistake, for several reasons:

(i) many commercial packages present a wide range of possible techniques and options. It may well not be obvious to the archaeologist which technique(s) (if any) is (are) appropriate, and in desperation they may choose 'all'options, and create a mountain of output, even for a single task,

(ii) since software has been written with 'archaeologically sensible' options and defaults, or for specifically archaeological problems (e.g. *iastats, mv-arch, arcospace*), it would be a mistake not to at least consider it,

(iii) archaeologists should always be aware of the models and assumptions that underlie any technique, and should be prepared to assess whether a technique is appropriate for answering the question they are asking,

(iv) conversely, archaeologists need to be sure that their data are appropriate for the application of a particular technique (e.g. they should not treat 'labels' in the same way as variables),

(v) archaeologists need to be able to recognise the occasions when the software produces bizarre results, possibly due to (iii) or (iv) above, or any other cause.

The subject of classification is not a fixed body of knowledge, but is continually growing through both pure and applied research. Many important developments are signalled in (for example) the *Journal of Classification*, which can be a useful source of ideas. There is no reason why archaeologists should not develop their own techniques to deal with their own particular problems, but they need to be aware of work in other disciplines, to avoid wasting time 're-inventing the wheel'. They also need to examine carefully the questions they are asking, to make sure that their choice of technique is appropriate.

Finally, we must ask whether classification is as relevant now, in an age of post-processual archaeology, as it was in previous phases of archaeological development. Certainly, archaeological theory and aims have changed – we no longer seek for the detailed 'evolutionary' progression of artefact types, or possess the optimism of the 'New Archaeology', that the objectivity of 'automatic classification' would rapidly surpass our human endeavour. Now, we see a more multifaceted approach to archaeology, with each archaeologist 'doing their own thing'. What does this imply for the theory and practice of classification?

There is the practical point that archaeologists gather their data from a wide variety of sources – their own fieldwork, as well as published and unpublished work of other archaeologists. Whatever the questions being asked of the data, there must be a common terminology, or at least a lowest common denominator, across the sources being used. Otherwise, the study will be of the differences between archaeologists instead of, or confounded with, genuine patterns in the data. It is rarely possible to re-examine all the earlier material, and to re-cast it into one's current terminology (even if the material has actually survived). There is a conflict here, between the need to 'speak the same language', and the need to prevent the development of a subject from being fossilised by a terminology that it is impossible to change. One solution may be the use of 'dynamic classification' (ANDRESEN, MADSEN, this volume).

This brings us to the final point. Whether archaeologists are working solely on their own material, or on material from a variety of sources, there is still a need for 'good' classification. Clearly the aims of a classification (whether an explicit exercise, or implicit in the use of a particular terminology) are set by archaeologists, according to their needs, but the criteria against which the outcome must be judged, i.e. how well it meets those needs, come from within classification as a discipline in its own right. Archaeologists set the research agenda, but to have any chance of success they must follow legitimate procedures as imposed by the tools (from other disciplines) that they choose to use.

CLIVE ORTON
Institute of Archaeology
University College, London

Acknowledgement

I am grateful to Ash Rennie for reading and commenting on a draft of this paper.

BIBLIOGRAPHY

ALDENDERFER M.S. (ed.) 1987, *Quantitative Research in Archaeology*, Newbury Park, Sage Publications.

BAKER K.G. 1988, *Towards an archaeological methodology for expert systems*, in C.L.N. RUGGLES, S.P.Q. RAHTZ (eds.), *Computer Applications and Quantitative Methods in Archaeology 1987*, British Archaeological Reports International Series 393, 229-236.

BAYLISS A., ORTON C. 1994, *Strategic consideration in dating, or "How many dates do I need?"*, «Bulletin of the Institute of Archaeology», 31, 151-167.

BAXTER M.S. 1994, *Exploratory Multivariate Analysis in Archaeology*, Edinburgh, Edinburgh University Press.

BUCK C.E., LITTON C.D. 1991, *A Bayes approach to some archaeological problems*, in K. LOCKYEAR, S. RAHTZ (eds.), *Computer Applications and Quantitative Methods in Archaeology 1990*, British Archaeological Reports International Series 565, 93-9.

CHANG K.C. 1967, *Rethinking Archaeology*, New York, Random House.

CORMACK R.M. 1971, *A review of classification*, «Journal of the Royal Statistical Society Series A», 134, 321-353.

DALLAS C. 1992, *Relational description, similarity and classification of complex archaeological entities*, in G. Lock, J. MOFFETT (eds.), *Computer Applications and Quantitative Methods in Archaeology 1991*, British Archaeological Reports International Series, 577, 167-78.

DAVIES B.J., RICHARDSON B., TOMBER R.S. 1994, *The archaeology of Roman London, Volume 5: A dated corpus of early Roman pottery from the City of London*, Council for British Archaeology Research Report 98.

DE VORE I. 1968, *Comments*, in L.R. BINFORD, S.R. BINFORD (eds.), *New Perspectives in Archaeology*, New York, Aldine.

DORAN J.E., HODSON F.R. 1975, *Mathematics and Computers in Archaeology*, Edinburgh, Edinburgh University Press.

DRAGENDORFF H. 1895, *Terra sigillata*, «Bonner Jahrbücher», 96, 18-155.

DROWER M.S. 1985, *Flinders Petrie: A Life in Archaeology*, London, Gollancz.

DURHAM P., LEWIS P., SHENNAN S. 1995, *Artefact matching and retrieval using the Generalised Hough Transform*, in J. WILCOCK, K. LOCKYEAR (eds.), *Computer Applications and Quantitative Methods in Archaeology 1993*, British Archaeological Reports International Series 598, 25-30.

EVERITT B.S. 1980, *Cluster Analysis,* (2nd edition), London, Heinemann.

GERO J., MAZZULLO J. 1984, *Analysis of artifact shapes using Fourier series in closed form*, «Journal of Field Archaeology», 11, 315-22.

GILLAM. J.P. 1970, *Types of Roman Coarse Pottery Vessels in Roman Britain*, 3rd edition.

GORDON A.D. 1987, *A review of heirarchical classification*, «Journal of the Royal Statistical Society Series A», 150, 119-137.

GRASLUND B. 1987, *The Birth of Prehistoric Chronology: Dating methods and dating systems in nineteenth-century Scandinavian archaeology*, Cambridge, Cambridge University Press.

HAGSTRUM M.B., HILDEBRAND J.A. 1990, *The two-curve method for reconstructing ceramic morphology*, «American Antiquity», 55(2), 388-403.

HALL N.S., LAFLIN S. 1984, *A computer aided design technique for pottery profiles, database*, in S. LAFLIN (ed.), *Computer Applications in Archaeology 1984*, Birmingham, University of Birmingham Computer Centre, 178-88.

HARDY-SMITH A. 1974, *Post-medieval pot shapes: a quantitative analysis*, «Science and Archaeology», 11, 4-15.

HODDER I., ORTON C. 1976, *Spatial Analysis in Archaeology*, Cambridge, Cambridge University Press.

HODSON F.R. 1970, *Cluster analysis and archaeology: some new applications and developments*, «World Archaeology» 1(3), 299-320.

HODSON F.R., SNEATH P.H.A., DORAN J.E. 1966, *Some experiments in the numerical analysis of archaeological data*, «Biometrika», 53, 311.

KRIEGER A.D. 1944, *The typological concept,* «American Antiquity», 9, 271-88.

LEESE M.N., MAIN P.L. 1983, *An approach to the assessment of artefact dimension as descriptors of shape*, in J.G.B. HAIGH (ed.), *Computer Applications in Archaeology 1983*, Bradford, School of Archaeological Sciences, University of Bradford, 171-80.

LIMING G., HONGJIE L., WILCOCK J.D. 1989, *The analysis of ancient Chinese pottery and porcelain shapes*, in S.P.Q. RAHTZ, J.D. RICHARDS (eds.), *Computer Applications and Quantitative Methods in Archaeology 1989*, British Archaeological Reports International Series, 548, 363-74.

MAIN P.L. 1981, *A method for the computer storage and comparison of the outline shapes of archaeological artefacts*, unpublished Ph.D thesis, London, Council for National Academic Awards.

MAIN P.L. 1986, *Accessing outline shape information efficiently within a large database*, in S. LAFLIN (ed.), *Computer Applications in Archaeology 1986*, Birmingham, University of Birmingham Computer Centre, 73-82.

MONTELIUS O. 1903, *Die typologische Methode: Die altere Kulturperioden im Orient und in Europe*, Stockholm, Selbstverlag.

ORTON C.R., TYERS P.A. 1992, *Counting broken objects: the statistics of ceramic assemblages*, in A.M. POLLARD (ed.), *New Developments in Archaeological Science*, Oxford, Oxford University Press.

PETRIE W.M.F. 1904, *Methods and Aims in Archaeology*, London, Macmillan.

PITT-RIVERS A.H.L.F. 1906, *The principles of classification*, in J.L. MYERS (ed.), *The Evolution of Culture and Other Essays*, Oxford, Clarendon Press.

PLIQUE A.E. 1887, *Etude de céramique arverno-romaine*, Caen.

RHODES M. 1979, *Methods of cataloguing pottery in Inner London: an historical outline*, «Medieval Ceramics» 3, 81-108.

RICHARDS J.D. 1987, *The significance of form and decoration of Anglo-Saxon cremation urns*, British Archaeological Reports, British Series 166, Oxford.

ROUSE I. 1960, *The classification of artefacts in archaeology*, «American Antiquity», 25, 213-23.

SHEPARD A.O. 1956, *Ceramics for the Archaeologist*, Washington, Carnegie Institute of Washington.

SMITH C. ROACH 1854, *Catalogue of the Museum of London Antiquities collected by, and property of, Charles Roach Smith*, London, privately printed.

SMITH C. ROACH 1859, *Illustrations of Roman London*, London, privately printed.

SOKAL R.R., SNEATH P.H.A. 1963, *Principles of Numerical Taxonomy*, San Francisco, Freeman.

TUKEY J.W. 1977, *Exploratory Data Analysis, Reading*, Mass., Addison-Wesley.

TYLDESLEY J.A., JOHNSON J.G., SNAPE S.R. 1985, *"Shape" in archaeological artefacts: two case studies using a new analytic method*, «Oxford Journal of Archaeology», 4(1), 19-30.

WEBSTER G. (ed.) 1964, *Romano-British coarse Pottery: a students guide*, Council for British Archaeology Research Report 6, London, Council for British Archaeology.

WILCOCK J.D., SHENNAN S. 1975, *Computer analysis of pottery shapes*, in S. LAFLIN (ed.), *Computer Applications in Archaeology 1975*, Birmingham, University of Birmingham Computer Centre, 98-106.

WRIGHT R. 1989, *Doing Multivariate Archaeology and Prehistory: handling large data sets with MV-ARCH*, Sydney, Dept. of Anthropology, University of Sydney.

ABSTRACT

Classification has been an important archaeological activity for at least a century. It should not be seen as an end in its own right, but as a tool that enables archaeologist to compare and communicate. It can also be seen as mathematical activity, the study of relationships between entities in a multi-dimensional space. Comparison of these approaches leads to a set of criteria for a "good" classification. A wide and growing range of techniques is available, but more fundamental issues such as the choice of variables and the level of analysis must also be considered. The history of classification in archaeology shows a period of optimism followed by one of disappointment; the recent development of techniques more suited to archaeological needs may enable a middle view to be taken. Finally, the role of classification is assessed in the light of current trends in archaeological theory.