7

The computer representation of space in urban archaeology

J. W. Huggett
(Department of Archaeology, University of Glasgow, Glasgow G12 8QQ)
M. A. Cooper
(Archaeology Section, Hereford and Worcester County Council, Tetbury Drive, Warndon, Worcester WR4 9LS)

7.1 Introduction

The depth of deposits and wide-ranging assemblages of artefactual and environmental material on deeply stratified urban archaeological sites attest to the frequent modification of the physical landscape through a wide variety of activities over time. However, the quantity and complexity of these deposits and assemblages together with frequent post-depositional disturbance means that the problems of defining and understanding the use of physical space on such sites can be at their most extreme. The stratigraphic matrix is of central importance in the structuring of the investigation of site morphology — the matrix may be used to combine entities into structure and phase plans which form the basis for a discussion of the spatial development of the site. In this way, the matrix, essentially a time rather than space dimension, structures the analysis of space.

The rôle of computers in the representation of space in urban archaeology has had a somewhat chequered history. This is reflected in the fact that, while a number of computer-based methods for examining stratigraphic sequences have been devised, none have gained widespread acceptance. Computer programs have exclusively concentrated on the problem of sorting and representing matrix data — the analytical aspect has received much less attention. With few exceptions, such as Bryan Alvey's HINDSITE system (Alvey 1989), there is a yawning software gap between the generation of a sequence matrix and the creation of structure and phase plans.

It is worth investigating why computers have had such a limited impact. On the face of it, computers should be eminently suitable for application to stratigraphic matrices at least: these are created using well-established techniques which suggests that algorithms could be developed which could be computerised. Indeed, this is the area in which most efforts at computerisation have concentrated (for example, Cooper 1987, Wilcock 1975, Haigh 1985, Rains 1985, Ryan 1989, Ryan 1988). In most cases, implementations have been restricted to the checking of the logical consistency of relationships and removal of links which are not necessary for the construction of the sequence. Production of a matrix diagram was generally still carried out by hand, using the list of links generated by the computer program. In Ryan's recent critique of computerised approaches to stratigraphy, two particular problems are isolated: first, the inability of computer software to handle large volumes of data, and secondly, the poor match between the functionality of the software and the archaeologist's requirements (Ryan 1989, p. 327).

The problem of the sheer volume of data and methods of facilitating the processing of such large bodies of data have certainly exercised the minds of those working on computer stratigraphic software. In the early days (ten years ago), various complex methods were used in an attempt to enable underpowered computer equipment with limited memory and storage facilities to cope with a reasonably-sized site. Today, with the falling cost of fast computers with high storage capacities, this is less of a problem, but there is still much discussion over the best way to represent stratigraphic information within the computer (e.g. Cooper 1987, Ryan 1988).

The functionality of the software remains an obstacle to its acceptance. Most authors have adopted a mainframe style of presentation, with instructions entered at the command line, output presented as error messages indicating logically invalid relationships, and a listing of the sequence. No interaction between computer and user occurs between the entry of the command and the resulting output. This 'black box' approach has clearly antagonised some archaeologists (e.g. Harris 1975). The opportunity to modify the sequence chain during its construction is not allowed for: consequently the computer has to either complete a full pass through the data set each time before alterations can be made or drops out of the program each time an error is encountered so that a typical session consists of interminable edits and re-runs. Furthermore, such "computer criticism" of a stratigraphic matrix is aggravated by the lack of any indication of the methods used to generate the conclusions or justification of the results — the computer blindly applies the logical procedures defined by the programmer. The idea that a computer might be daily checking the consistency of relationships almost as they are revealed has also caused difficulties on site (McVicar & Stoddart 1986).

Attempts have been made to improve the functionality of software. Ryan (1988a), for example, has demonstrated his GNET system which lays out a graph using any data (not necessarily stratigraphic) as long as it consists of nodes and edges. The resulting diagram can then be directly modified by the user, links edited, and a choice of representation selected — all links or only those necessary to the sequence. The software employs a graphical user interface, making it intuitive to use and enhancements have included the ability to query the stratigraphic database using the matrix diagram. A different approach has been adopted by Alvey with his HINDSITE software (Alvey 1989). Here, a commercial computer package has been modified for archaeological use and a quasi three-dimensional model is constructed from two-dimensional context plans with the stratigraphic relationships providing the third dimension. In this case, the computer is actually being used in a way that cannot be duplicated by hand, rather than performing a standardised procedure faster than would otherwise be possible. The developments in three-dimensional CAD software such as AUTOCAD Release 10 means that three dimensional construction and manipulation of matrices is now feasible, although its utility in this regard remains to be demonstrated.

Computer-based approaches to spatial representation have therefore tended to follow a well-trodden course — the validation of sequences and the creation of Harris matrices, with more recent work concentrating increasingly on the graphical representation of the resulting model. In many cases, the generation and manipulation of the matrix diagram appears to be seen as the end of the process. It is unfortunate that not only does the existing software fail to measure up to expectations, but it is limited in its methodology and application. Software can give substantial assistance in the creation of a matrix, but the methodology used is built around only one form of representation and what is ultimately lacking is computer-assisted analysis of the resulting representation.

7.2 Alternative representations

Perhaps the most obvious alternative to the two-dimensional Harris matrix is its conversion into a three-dimensional model using computer software. The advantage of this would be that both temporal and spatial relationships could be represented in the one model. To some extent, however, three-dimensional representations are symptoms of what is increasingly becoming recognised in the computing world as the 'because we could' syndrome (e.g. Seymour 1989) — because three-dimensional modelling is possible, it is therefore a good method to use, regardless of its actual utility. The belief in some archaeological quarters seems to be that the three-dimensional computer representation of site matrices will remove the need for a methodologically sound means of two-dimensional representation (Daniels 1989), failing to recognise that such models remain founded in the same two-dimensional methodology. The interpretability of such constructs is taken for granted, as is the ability of an average computer to process such data. Whether these developments are of actual benefit to the archaeologist is open to debate. Some would suggest that computer-assisted methods should be restricted to those which are already practised, rather than define new, computer-driven, procedures (e.g. Ryan 1989). While this would appear to limit computer-based research, archaeological software application developers have to resist the temptation of diverting their resources into products that are fun to build but are neither particularly useful nor usable. Three-dimensional diagrams may be best reserved for the presentation of data rather than the analysis of it, at least in terms of the site matrix.

Another representation is already in use, but a computer-based version has not yet been developed. The 'Carver Matrix' employs a slightly different approach by explicitly combining entities which make up an archaeological site into higher order analytical units (Carver 1979), a procedure which is normally carried out as part of the post-excavation process regardless of the representation method used. A hierarchical method is adopted, in which contexts are the lowest order unit, features consist of groups of contexts, and structures are made up of features. Contexts which are not identified as belonging to features are called the 'featureless set.' In the diagrammatic representation of this method, the featureless set of contexts are drawn as horizontal lines and provide a matrix into which the higher order entities are fitted. Features are represented as vertical arrows with the lower point beginning at the latest context which it cuts or seals, and the higher point finishing at the first context which seals the feature. In terms of the representational advantages, Carver's method is likely to result in a diagram which more closely models the layout of the site since it separates out those entities which form major surfaces and uses them as a framework for the matrix. The reduction involved in the definition of features results in a simpler and thus clearer diagram, and the relative lives of entities are more easily assessed. Viewed as a representation which is one step removed from the Harris matrix and thus one step nearer to a representation of the physical landscape, the method has much to commend it, particularly if both methods are used in parallel.

The development of a computerised Carver matrix system could therefore have much to offer and would build upon the existing computer-based Harris matrix systems. The Harris data could be used to create a set of features, automatically or otherwise, and the interpretative diagram created directly. Indeed, this would reproduce the process of the grouping of entities and simplification of stratigraphic units which is commonly conducted as part of the post-excavation analysis (e.g. Museum of London 1986, Appendix 5).

A different approach would be to attempt to solve some of the problems which arise during the analysis of whichever representation is used. Indeed, it may be suggested that the need to develop analytical software should be of greater priority than the enhancement of a particular representational method — solving problems which are encountered now may be seen as being of greater benefit. One such area which may be usefully tackled is the correlation of sequences.

7.3 Correlating sequences

Although the stratigraphic matrix provides a basis for the reconstruction of the past landscape, it is not an ideal or perfect tool. Not only are there the problems commonly associated with Harris matrices — the crossovers resulting from the compression of the data into a two-dimensional form, and the tendency for single and complex related entities to be exploded across the diagram, for instance (see Cooper 1989) — but the depth and disturbance common in urban sequences may cause major difficulties in relating stratigraphic entities. For example, the reconstruction of site morphology on deeply stratified sites may be seriously hampered where the stratigraphic sequence divides into a number of distinct sequences bearing no direct stratigraphic relationship with each other. This may arise where sequences build up on either side of a wall, or where a cellar has been inserted, or where the same ground surface is exposed for a long period and subject to many deliberate disturbances, many of which do not intersect, or where later terracing has removed the tops of negative impacts together with their original stratigraphic relationships.

Since the accurate reconstruction of site morphology and human activity is likely to rely on the ability to assess chronological relationships between all deposits, whether stratigraphically related or not, the sequences need to be correlated in some way. Factors such as similarity of constituents, shape, and spatial relationships may be used to link otherwise stratigraphically unrelated entities, but where such attributes are not applicable, the artefactual assemblages contained within the entities are frequently used.

Clearly, using artefactual assemblages in this way may give rise to additional problems, for while the artefacts from an entity ought to give a broad indication of its date, it necessarily implies that the entity should receive all its artefactual components at the time of its creation and use, and no inclusions should be received subsequently. Furthermore, not all entities will contain artefactual material, or artefactual material that is datable. Most important, particularly on urban sites, is the problem of 'derived' (residual or intrusive) material — artefacts which are not in their original context of deposition whether as a result of human activity or post-depositional disturbance by non-human agencies. Four types of assemblage may be defined (Cooper 1989):

Type 1: Primary Assemblages, containing material contemporary with the entity's creation and use.
Type 2: Residual Assemblages, containing material dating to an earlier period than the entity's creation and use.
Type 3: Intrusive Assemblages, containing material dated to a later period than the entity's creation and use.
Type 4: Mixed Assemblages, containing material from a combination of two or more of the above types.

7.4 Urban pottery seriation

One of the only detailed discussions of the use of derived material is that by Carver in his paper on urban pottery seriation (Carver 1985). In this procedure, pottery types are chronologically ordered on the X-axis of the diagram with the earliest on the left and the latest on the right. The stratigraphic entities are arranged on the Y-axis with the earliest at the bottom and the latest at the top. The incidence of pottery is drawn out using this structure, and where the entities are not stratigraphically related, the pottery assemblage may be used to re-order and group the entities, with those containing later material being moved upwards and those with no late pottery moving down (Carver 1985, p. 357–9). According to Carver, this method can be taken further, with the shape of the leading edge of the diagram being used to draw conclusions about the site's occupational history. Thus, within strict limits, Carver argues that the diagonal of the diagram can reveal information about the overall character of the site and the way in which it has been exploited, allowing the identification of periods of gradual growth (slope), dense contemporary occupation (cliffing), dumping (indentation), and levelling (plateaux) (Carver 1985, p. 360–1). Whether this level of interpretation is justified is another matter, but the basic principle can still provide a valuable tool for manipulating sequence and artefact data in combination.

The extent to which this method can be successfully applied is clearly dependent upon the explicit definition of the nature of each assemblage used, which is in itself a complex task. However, the method can enable the repositioning of floating chains of entities and lead to the redrawing of the matrix (under both the Harris and Carver methods) providing that this does not conflict with existing recorded stratigraphic relationships.

7.5 A computer-based approach

The correlation of sequences is an area which is particularly applicable to a computer-based approach. This builds on what has gone before, but also attempts to assist with the problems encountered in the construction of stratigraphic sequences, particularly in urban contexts. In addition, it is anticipated that the process of constructing computer software to carry out this function, many of the concepts and assumptions inherent in the approach may be tested and re-assessed.

It is proposed to develop a system which will correlate stratigraphic entities based primarily upon the seriation methods described by Carver (1985) although clearly other methods are available which might prove more suitable. The initial system, which is currently under development, will simply take an ordered file of contexts and a chronologically ordered file of artefact types and produce a diagram showing the incidence of such artefacts in contexts. Where the contexts are not stratigraphically related, it would then be possible to re-order those contexts on the basis of the artefacts within them. Initially, this will be a relatively crude system, performing little more than the production of a graph, although this alone is considered to be an extremely valuable and time-saving device.

Future development will proceed through a series of defined stages. First, it should be possible to interactively re-order the context and/or artefact axes from the diagram, rather than have to modify the data files directly. Secondly, the computer ought to be able to check the modifications made by the operator: for instance, an attempt to move a context higher up the sequence than is possible given its stratigraphic relations should generate an error. Thirdly, the ability to interrogate the database through the diagram would be added, so that the diagram is not created in the abstract but the nature and type of each of the elements of the data can be recalled. Fourthly, it might be useful to be able to conduct "what-if" experiments, modifying doubtful links in the sequence or typology and observing the effect on the re-drawn diagram.

Ultimately, it is intended to build in a degree of artificial intelligence, so that the diagram is automatically generated and the "best-fit" solution in terms of both the sequence and the typology derived according to a series of established rules. It should be possible for the user to query the results and for the computer to justify decisions, in order to avert the 'black box' syndrome. The system would then consist of a matrix generation facility and a seriation facility. It would seem reasonable therefore, in terms of the archaeological utility of the product, to include an interface to a graphics system to enable the creation of structure and phase plans from the results.

Approaching the problem in this way means that at each stage, the resulting software will be a usable tool, and that the theory and assumptions implicit in the methodology will be made explicit and testable. In the end, the aim is to provide the excavator with a tool which enables stratigraphic entities to be ordered, subsequent floating chains correlated, and the appropriate phase plans generated.

Bibliography

ALVEY, B. 1989. "Hindsite", *Archaeological Computing Newsletter*, 19: 4–5.

CAA 1985. *Computer Applications in Archaeology 1985*, London. Institute of Archaeology, University of London.

CARVER, M. O. H. 1979. "Notes on Some General Principles for the Analysis of Excavated Data", *Science and Archaeology*, 21: 3–14.

CARVER, M. O. H. 1985. "Theory and Practice in Urban Pottery Seriation", *Journal of Archaeological Science*, 12: 353–366.

COOPER, M. A. 1987. "Using Archaeological Stratigraphy", *Archaeological Computing Newsletter*, 12: 3–13.

COOPER, M. A. 1989. "The Definition and Representation of Space in Urban Archaeology", *Scottish Archaeological Review*, 6: 75–88.

DANIELS, R. 1989. "The Definition and Representation of Space in Urban Archaeology: Some Comments", *Scottish Archaeological Review*, 6: 88–89.

HAIGH, J. G. B. 1985. "The Harris Matrix as a Partially Ordered Set", *in* CAA 1985, pp. 81–90.

HARRIS, E. C. 1975. "The Stratigraphic Sequence: A Question of Time", *World Archaeology*, 7: 109–121.

MCVICAR, J. & S. STODDART 1986. "Computerising an Archaeological Excavation: The Human Factors", *in Computer Applications in Archaeology 1986*, pp. 225–227, Birmingham. University of Birmingham.

MUSEUM OF LONDON 1986. "Department of Urban Archaeology Archive Report Writing Manual".

RAINS, M. J. 1985. "Home Computers in Archaeology", *in* CAA 1985, pp. 15–26.

RYAN, N. S. 1988. "Data Structures for Stratigraphic Analysis", *Archaeological Computing Newsletter*, 14: 1–11.

RYAN, N. S. 1989. "Browsing through the Stratigraphic Record", *in* Rahtz, S. P. Q. & Richards, J. D., (eds.), *Computer Applications and Quantitative Methods in Archaeology 1989*, International Series 548, pp. 327–334. British Archaeological Reports, Oxford.

SEYMOUR, J. 1989. "The 'Because We Could' Syndrome", *PC Magazine*, 9 (3): 79–80.

WILCOCK, J. D. 1975. "Archaeological Context Sorting by Computer", *in Computer Applications in Archaeology 1975*, pp. 93–97, Birmingham. University of Birmingham.