Thirty years of computer archaeology and the future, or looking backwards and forwards at the same time while trying not to twist one's neck.

Irwin Scollar, Laboratory for Field Archaeology, Rheinisches Landesmuseum, Colmantstr. 14, D5300 Bonn 1
West Germany

The title of this paper is wrong. It really should read something like "25 or so years", which makes things a nice round quarter-century. I believe that the paper on seriation published by Robinson in 1951 did not immediately result in a computer program. Nonetheless it can be regarded as the beginning of the application of a particular class of numerical methods in archaeology. Elementary statistical methods surely were earlier, but their first use is lost in the mists of antiquity. Robinson's paper was one of the first attempts at direct evaluation of formal archaeological data (types etc.) culled and classified by hand, but sorted by a numerical method. Site data was probably first put into a real machine sometime around 1959 by Ihm and Gardin according to Cowgill (1968) and Ihm gives a very incomplete reference to this in Ihm(1978). Such is the conventional view.

However, there are two other classes of archaeological data which have been frequently subject to computer treatment, one the cataloging of information (databanks), the other the processing of information obtained from one of the ancillary sciences (dating, prospecting, etc.). Although machine treatment of catalog data using mechanical punch card systems also goes back to the epoch before computers (Gardin 1956), probably the latter class has the honor of being the first data ever treated by machine (electronic). During the 1950's at Cambridge, Robert Cook embarked on an ambitious program of sampling kilns for magnetic dating. John Belshe, then of the Department of Geodesy and Geophysics worked as his assistant and made the measurements on the samples. He has told me that he wrote what I believe to be the first program ever to be applied to archaeological data to evaluate the astatic magnetometer readings for the EDSAC computer and used it in 1957. Unfortunately the results from much of this work were not completely published, and nothing was ever mentioned about the data evaluation as far as I can determine.

I have been unable to determine when the first data banks were set up to deal with archaeological museum records since the intermixture with earlier mechanical techniques is too great. Probably the SELGEM group at the Smithsonian can claim a first, but I would like to hear about other contenders.

Therefore I think that we should consider the development of computer methods in archaeology within the outlines of three more or less distinct classes of problems, because the aims, methods and quantity of data treated have been different in each case. I consider the minor class of problems in metrology or elementary site recording statistics as belonging to the category of data analysis. The three categories, processing data about archaeological entities, catalogs or the like of those entities, and measurements made in connection with them via other disciplines appear to exhaust the possibilities. At least I do not know of any others.

The three categories have been approached by quite different kinds of people as a rule. The first by archaeologists and statisticians, the second by museum people who may or may not be archaeologists, and the third by physicists or engineers. It is not too surprising therefore, that the "hard" scientists were first. The early computers were usually located in places where physical scientists had access to them, and they were designed for number and not for text crunching (nibbling actually). The data bank class, given the means of the first and second generation with punched card and paper tape data entry, automatically forced attention to problems in compact coding of information in a near-numerical form. Search and storage strategies followed. The literature up to about the early 1970's is full of different coding schemes. The first category usually left the major problem of type coding to the archaeologist to do by hand and work was concentrated on statistical approaches for dealing with the pre-classified data, either to improve its ordering or refine the classification. Somewhat later, geographical context was introduced and studies in spatial ordering appeared in the late 1960's and 1970's.

I believe that the developments in all three categories are technology driven, that is, advances in hardware, operating software and high level languages offer opportunities for invention which drive progress. This progress has usually taken place outside the archaeological application sphere in related subjects. People with training in several subjects have been able to take advantage of progress elsewhere. For example, progress in numerical analysis which was stimulated by the more efficient machines of the second generation resulted to a transfer of methods to geophysics, and then ultimately to archaeological prospecting. I myself was greatly influenced by the paper which appeared in 1958 by Dean (1958) to try digital filtering methods on resistivity data and published in 1959, although the plotted results were in fact hand drawn from the calculated values. It wasn't until a plotter became available on a second generation machine in 1963 that plots were made automatically (Scollar and Krückeberg 1966). But a punched paper tape system for field readings from our home made magnetometer was purchased in 1961. By the time the hardware was ready to make tapes for the Standard Elektrik Lorenz ER 56 at Bonn University, that first generation machine (the only one ever made by IT&T) was taken out of service and replaced by the IBM 1410 which was actually used for some calculation.

In this rambling review, it is probably better to carry through category by category, even if this means some jumping back and forth in time. I envisage the hierarchy from the Olympus of data analysis, down to the earthly work of cataloging and data banks, and to the underground activities of the basic measurement evaluation gang in the cellar. I do not pretend to completeness and will be grateful to any reader who will write to me and tell me about what I have forgotten. If I have omitted your favorite paper, it is either through ignorance or choice and I leave it as an exercise for the reader to decide which applies.

DATA ANALYSIS

By data analysis, I mean procedures for classifying, ordering and grouping of data from archaeological objects, collections of objects or sites, relative to characteristics determined by empirical need or observation, or perhaps as an attempt to derive the needed characteristics themselves from some "raw" data. Hodson and Doran (1975) have called this the "Mainstream", although I believe that far less data has been usefully processed at this level than in the other two categories of this study. Probably, because of its intrinsic intellectual appeal, data analysis has received the largest coverage in the literature, although data bank studies being by nature bulkier take up a greater number of centimeters on my bookshelf. In essence, it involves all of the difficulty of automatic pattern recognition and classification, also one of the subjects which has probably engendered the most hope and the most disappointment in the application of computer techniques.

That seriation, the ordering or sequencing of archaeological complexes, received attention first is not an accident. There were well established manual procedures for carrying out this routine operation on cemeteries, stratification, etc. long before computers were in existence. The first papers after Robinson (1951) were unsophisticated, but by the mid 1960's quite professional mathematical attention was being given to the problem as emphasized by the series of papers by Kendall and Wilkinson. Interestingly enough, one of the most effective algorithms for seriation in terms of computational simplicity and speed was developed by Goldmann,an archaeologist and programmed by Kammerer. Variants of this method are in use today (Goldmann 1972). Interest in seriation peaked with a series of papers in the early and mid-seventies. Since then little new has appeared and an apparent stability has been achieved, with the exception of recent work by Ihm. The more elementary techniques which are readily understood by archaeologists have retained favor. A simulation study done by Galloway, Graham and myself (1976) showed that most of the available methods produced very similar results, so that a choice of computationally simple ones is justified. However, almost all of the published methods depend on manipulation of large matrices of data with considerable waste of storage and time. The introduction of the microcomputer has now revived interest in the simple algorithms modified for space saving and fast treatment in order to handle large data sets in small machines. A current software project (low key) in my laboratory is the development of transportable software written in PASCAL for small machines, and making use of dynamic memory allocation. Unfortunately, there are really very few archaeological data sets which are suitably well linked and studied for seriation to be of much use. I can't think of more than a dozen examples of large data sets in the literature, of which one of the more notable is the work by Goldmann himself, or the recent publication of Perrin(1980) using Goldmann's algorithm. Experience with users of seriation programs has shown that the mere classification and typing of the raw data usually gives the archaeologist such good insight into the chronology that the ultimate machine work only confirms his views.

CLUSTERING & CLASSIFICATION

Classification and intuitive clustering have been the life's work of archaeologists almost from the very beginnings of the subject. The mathematical clustering algorithms which appeared during the 50's and 60's for use in the life sciences were soon applied to archaeological data in the mid 60's, mainly by Hodson who discusses merits and demerits in great detail in Doran and Hodson (1975). Since the early seventies, large well tested computer packages like CLUSTAN which offer most known algorithms have reduced the problem for the archaeologist to data entry. In some respects this is also true for elementary statistical techniques. Here the SPSS package appears to be the favorite of most users at university computing centers. To my knowledge, some of the newer clustering algorithms which have appeared in the specialized literature Jarvis (1977) for example, have not been used in archaeology. The use of well tested packages is to be encouraged because it allows users to concentrate on problems of data selection and entry, rather than on programming. On the other hand, it converts the apparent objectivity of computer methods into subjectivity at a new and obscure level, namely in the choice of algorithm. There seems to be little awareness on the part of many archaeological users of packaged programs of how dangerous the choice may be. Hodson has pointed out many of the pitfalls, but maybe more warning is desirable. Perhaps the packages should be marked with a warning like those on cigarette packages in some countries. In a package distributed from our laboratory in Bonn during the early 70's we included the warning on the distribution tape. Since no one ever seemed to make use of the package, little harm was done anyway.

In the future, the cheap color graphics available on small machines ought to help a lot in clustering and classification. It will be necessary to adapt the algorithms to the reduced memory space available at present, but this limitation ought to disappear by the end of the decade.

Incorporation of spatial or geographical information in data analysis has been a slow development. Most computing centers had plotters by the mid-1960's, and techniques were converted to using this kind of hardware (Upham 1976).

Standard mapping and contouring packages such as SYMAP were used by the early 1970's and are now very common. On-line color displays, available in some places by the mid-70's are now in wider use. There are no publications which describe results making full use of their capabilities, but this should come about very shortly. The smaller micros with their direct attachment to a home color TV will surely encourage the trend, despite current lack of good resolution. As falling memory prices make higher resolution possible, the trend should accelerate. Subtle algorithms derived from advanced computer graphics will be used for data display. The analysis will need methods beyond thosedescribed in Hodder and Orton (1976) to make full use of the new capabilities. Probably,spatial analysis is the area where we can expect the most interesting development in the near future.

An offshoot of early classification efforts coupled with information retrieval evoked attempts at automated drawing of objects, Gardin (1971). Prior to computers there were and still are many attempts to use pantographs or camera obscuras for aid in drawing archaeological objects. Recently, television cameras have been coupled to small micros for the same purpose. None of the techniques published in the literature appear to have achieved wide use, probably because the percieved character of shapes of archaeological objects contains far too many variables to be seized by simply following the curves of an outline. It is likely that in the future, this will continue to be a problem and that automated drawing methods for complex three dimensional objects will fail. Hopes for their success are intimately linked with the problems of data entry to data base systems, since tedious coding of shape descriptors has proven to be the death-blow to many of the elaborate information retrieval systems proposed in the euphoria of the period of the second computer generation.

DATA BANK SYSTEMS

I think that more effort has been devoted to producing systems which will allow archaeologists to find objects of interest in museum collections than any other single endeavor. Probably more systems of this kind have failed than have succeded, usually over the high hurdle of data entry whose costs in relation to results were seriously understimated at system design time. The schemes of the punch card age were never used for really large masses of archaeological objects. One impressive early effort, the mechanical file of tools from the middle east (Gardin 1956) was worked on for many years. At a public meeting in France it was recently stated that no one ever consulted the data afterwards. If true, this is an interesting comment on some systems. Many of the early approaches required coding in order to format data and allow search in a reasonable time.

With the advent of the more powerful machines of the 3rd generation in the 1960's, including time sharing and text oriented processing, a new phase was reached. A few working systems were produced which are around even now. The development of smaller machines in the late 70's led to the rise of the commercial data base management system which can easily be adapted to archaeological use in maintaining local inventories. In the last few years, relational data base approaches have simplified things, even though they require a difficult standardization of terminology. The current tendency is away from big bases and toward smaller specialized indices which encourage data entry by the users. The larger hierarchic structures using main frames continue however. Both approaches will probably exist side by side for quite a time, although ease of use, low price and reliability will tend to direct most new projects to the micros and lower level minis. The flurry of coding problems of the 1960's has now died out, since it was quite quickly discovered that no one was willing to do the coding. The newer methods using standardized natural language, with the machine doing the coding internally if at all, appear to have more chance. At my laboratory we run several data banks side by side on our machine. One of them is the direct successor of a mainframe 1960'ish approach, the others are all based on a commercial system purchased as a software package in 1980. With the newer system, ease of data entry is high, search speed is reasonable if the data base is not too large, space is not excessive, and there is no programming of any significance required. Current developments are extending this technique to all but the very smallest micros. In the future, it will be quite unnecessary to write programs for data banks. The commercial packages, perhaps with some validation and forms management front ends (also bought), will replace the hand crafted products of the 60's and 70's. The smaller relational approaches concentrate attention on essentials--namely dimensioning the system to questions frequently asked of it, rather than to questions which might concievably be asked.

ANCILLARY SCIENCES

By ancillary sciences I mean the techniques which support prospecting, dating and other things in archaeometry. These are direct adaptations of mathematical methods usually developed in related sciences. As mentioned in the introduction, here were the first completely successful applications of computer techniques in archaeology. They were also mostly non-controversial at least among the archaeological customers. Modern archaeological prospecting is now almost unthinkable without computer analysis. This was not the general opinion when I produced a computer calculated and filtered resistivity map in 1959, and subsequent advances in automatic recording of magnetometer data were poo-poohed until the poo-poohers tried the techniques on very large sites. Computer evaluation of dating results were, as mentioned, common post-1957, and much effort also went into radiocarbon calibration curves with various smoothing techniques in the mid-1970's. In tree-ring dating , introduction of correlation techniques were first used with off-line measurements, but now on-line real time measurement with a laboratory micro is a commonplace.

Aerial archaeology provided the impetus for interactive analytical photogrammetry of oblique images. The first attempt was my paper published in 1975, followed by Palmer (1976) and others. On-line photogrammetry with simple analytical approaches were used with minis and micros in the late 70's. The first applications of space research developed image processing were also done at our laboratory in connection with the installation of a large image processing machine in the mid-1970's. Tests were made on machines in a commercial environment by me in 1971, but not published. Unfortunately, because of the complexity of the hardware and its cost, this work has remained unique until recently. In the future, the introduction of semiconductor image acquisition devices of high resolution, relatively cheap hardcopy gray scale or color reproduction devices and general cheapening of gray scale or color display devices should change all this. Up to now, applications have been mostly to archaeological air photos, x-ray images, ultraviolet and infra-red pictures of archaeological and art objects, and to thermal infra-red scanning of the prospecting type (Scollar 1977,1979). Magnetic and other types of prospecting have profited from the extended hardware available, and quite small cheap machines with color displays are beginning to be used for the purpose in the USA.

EXCHANGE OF DATA AND SOFTWARE

For all of the three categories discussed above, there has always been some interchange of software and data between workersactive in the field. Availability of commercial software packages has made this unnecessary for a number of problems. All pious wishes with regard to data and software interchange have never been achieved except in this commercial context over any significant period of time. However, only those products of wide interest outside archaeology can attract commercial software producers. Hence there will always be a small residue of applications which require hand-crafted solutions. Besides, one cannot have all the fun taken out of the business. It seems unlikely that the chaos of the past will be reduced in the hand crafted products, even with the introduction of highly structured high level languages. But maybe one will at least be able to read and follow someone else's program.

THE FUTURE?

I think that the great remaining area for progress is with visual data. Archaeology deals primarily with things seen in the first instance and it requires reduction of results with transformed visual output. However I don't expect that many will consult the literature on image processing and graphics--there are easily 15000 titles by now. Most people seem unaware of the resources available at a really good technical library and continue to re-invent hexagonal wheels which bump down the problem road as if nothing has ever been done before, at least to judge from some of the literature. Probably the most difficult problems in applications of computer methods to archaeology are knowing what has been done elsewhere in related conditions in other fields with success (and what has failed!), and knowing when to stop.

Selected bibliography-partly cited in text:

Data Analysis

J.C. Gardin, A.M. Richaud ed. Archéologie et Calculateurs. CNRS Paris 1970.
M. Borillo ed. Raisonment et méthodes mathematiques en archéologie, Ed. CNRS 1977
G.L. Cowgill, Computer applications in archaeology, Fall Joint Computer Converence AFIPS, 31, 1967 331-337.
G.L. Cowgill, Archaeological applications of factor, cluster and proximity analysis, American Antiquity, 33 1968, 367-375
J.E. Doran & F.R. Hodson, Mathematics and Computers in Archaeology, Edinburg, 1975
Les dossiers de l'archéolgie: l'analyse des objets archéologiques et les procédés statistiques d'interpretation, No. 42, Dijon, 1980
I. Graham, P. Galloway, I. Scollar, Model studies in computer seriation, Journal of Archaeological Science, 3, 1976, 1ff.
K. Goldmann, Zur Auswertung archäologischer Funde mit Hilfe von Computern, Die Kunde, 19, 1968, 1-8.
K. Goldmann, Zwei Methoden chronologischer Gruppierung, Acta Praehistorica et Archaeologica, 3, 1972,1-34
F.R.Hodson, P.H.A.Sneath, J.E. Doran, Some experiments in the numerical analysis of archaeological data, Biometrika, 53, 1966, 311-324
F.R. Hodson, D.G. Kendall, P. Tâtu, Mathematics in the archaeological and historical sciences, Edinburgh, 1971
F.R. Hodson, Searching for structure within multivariate archaeological data, World Archaeology, 1, 1969, 90-105
F.R. Hodson, Cluster analysis and archaeology, some new developments and applications, World Archaeology 1, 1970,299-320
F.Hole, M. Shaw, Computer analysis of chronological seriation, Rice University Studies 53, 1967
P. Ihm, Statistik in der Archäologie, Archaeo-Physika 9,Köln, 1978
R.A. Jarvis, Shared near neighbor maximal spanning trees for cluster analysis, Proc. 4th Int. Cof. Pattern Recognition, Kyoto, 1978 308-313 and in detail in Report TR-EE 77-45, School of Electrical Engineering, Purdue University, 1977
D.G. Kendall, A Statistical approach to Flinders Petrie's Sequence Dating, Bulletin of the I.S.I., 34th Session, Ottawa 1963, 657-680

D.G. Kendall, Incidence matrices, interval graphs and seriation in archaeology, Pacific Journal of Mathematics, 28, 1969, 565ff.
D.G. Kendall, Some problems and methods in statistical archaeology, World Archaeology 1, 1969, 68-76
D.G. Kendall, A mathematical approach to seriation. Phil. Trans. Roy. Soc. London, A 269 1970, 125-135
D.G. Kendall, Abundance matrices and seriation in archaeology, Zeitschrift f. Wahrscheinlichkeitstheorie, 17, 1971, 104-112
Les Méthodes Mathématiques de l'Archéologie, M. Borillo et. al Centre d'analyse documentaire pour l'archéologie, Marseille, 1971
C.Orton, Mathematics in archaeology, Collins, London 1980
P. Périn, R. Legoux, La datation des tombes mérovingiennes, Genève, Droz, 1980.
W.S. Robinson, A method for chronologically ordering archaeological deposits, American Antiquity, 16, 1951, 293-301
E.M. Wilkinson, Techniques of data analysis, seriation theory Archaeo-Physika 5, 1974 1-142

Shape Coding

J.C. Gardin, Le fichier mécanographique de l'outillage, Inst. Francais d'Archéologie de Beyrouth 1956 1-20
J.C. Gardin, Four Codes for Description of Artifacts, American Anthropologist 60, 1958, 335-357
J.C. Gardin, Etude d'un système integré d'acquisition et d'exploitation automatiques de données scientifiques, Rapport CNRS 36/1971/CRI-IRIA Projet III-43
J.M. Hollerbach, Hierarchical shape description by selection and modification of prototypes, MS Thesis, AI Laboratory, EE Dpt. MIT Jan. 24, 1975
O. Voss, Dokumentaionsproblemer Indenfor Arkaeologien, Kuml 1966, 97-134

Spatial Analysis & Graphics

N.I. Badler & V.R. Badler, SITE: A color computer graphics system for the display of archaeological sites and artifacts Report, U. Pennsylvania, Moore School of EE, Aug. 1977
I. Hodder & C. Orton, Spatial analysis in archaeology, Cambridge, 1976
C. Orton, Stochastic Process and archaeological mechanism in spatial analysis, Journal of Archaeological Science 9, 1982, 1-24
S. Upham ed. Computer graphics in archaeology, Ariz. State University Anthropological Research Paper 15, 1979

Data Base Systems

C.A.T.I.A. Une chaine automatisée de traitement de l'information archéologique, Rapport 1, 1973-5, Univ. Paris I, UER Art et d'archéologie.
R.G. Chenhall, Museum Cataloging in the Computer Age, Amer. Assoc. for State and Local History, Nashville, 1975
R. Ginouvès, A.M. Guimier-Sorbets, La constitution des données en archéologie classique, Ed. CNRS, Paris, 1978

J. Lemaitre et. al., Le Rida, Editions Adpf, Paris, 1980
A. Querrien, O. Buchsenschutz, J. Dorion, Carte Archéologique du Cher, Annales Litterires de l'Université de Besancon, 1979
J. Stewart, MDA Occasional Paper 4, Microcomputers in Archaeology, London, 1980
L.E. Vogler ed. The Arizona State Museum Archaeological Site Survey System, Tucson, 1980, Nr. 128

Dating

R.M. Cook & J.C. Belshé, Archaeomagnetism, Antiquity 32, 1958, 167
R.M. Clark, A Calibration curve for radiocarbon dates, Antiquity 49, 1975, 251ff.

Prospection, Air Photography, Image Processing

R. Palmer, Computer transcriptions from air photographs, Aerial Archaeology 2, 1978,5-9
I. Scollar, Einführung in die Widerstandsmessung, Bonner Jahrbücher, 159, 1959, 283-313
I. Scollar, F. Krückeberg, Computer treatment of magnetic measurements from archaeological sites, Archaeometry 9, 1966 61-71
I. Scollar, Automatic recording of magnetometer data in the field, Prospezioni Archeologiche 3, 1968, 105-110
I. Scollar, Fourier transform methods for the evaluation of magnetic maps, Prospezioni Archeologiche 5, 1970, 9-41
I. Scollar, Transformation of extreme oblique aerial photographs to maps or plans by conventional means or by computer. in Aerial Reconnaissance for Archaeology, CBA Research Report 12, 1975, 52-58
I. Scollar, L'informatique appliquée á la photographie aèrienne, L'archéologie 22, 1977 78-87
I. Scollar, B. Weidner, Computer production of orthophotos from single oblique images or from rotating mirror scanners, Aerial Archaeology 4, 1979, 17-28
A. Tabbagh, Deux nouvelles méthodes géophysiques de prospection archéologique, Thèse d'Etat, Paris 6, 1977
J.W. Weymouth, A magnetic survey of the Walth Bay site, Midwest Archeological Center, National Park Service, Lincoln, Nebraska, 1976
J. W. Weymouth, R. Nickel, A magnetometer survey of the Knife River indian villages. Plains Anthropologist 1977, 22-78.