Computer applications in archaeology*

by GEORGE L. COWGILL
Brandeis University
Waltham, Massachusetts

INTRODUCTION

In preparing this paper I have tried to give an accurate general picture of the kinds of things which have been done with computers by archaeologists and to give some of my own views about things which most need doing, but I have not tried to list every archaeological application of a computer ever made. The emphasis is mostly on work done in North America, and I believe that the coverage of significant work in this region is quite complete. I also include a good deal of what has been done in western and central Europe, but much less on eastern Europe and the Soviet Union. Gardin[1] has recently cited several important Russian publications which so far as I know are not available in English.

Archaeological use of computers is still in a very early stage. Several archaeologists had used machine-sorted punched cards much earlier, but the first archaeological applications of electronic data processing I know of were by Peter Ihm and by Gardin and Garelli in France around 1958 or 1959.[2] The earliest use I know of in this country was by James Deetz in 1960.[3] Subsequently there have been about 5 or 10 major projects involving computers and archaeology in North America, about the same number in the rest of the world, and many smaller ones. These projects have covered a wide range of quality and sophistication. As could be expected, in a few cases there was a too-hasty effort to use a machine because of the attraction of something so fashionable and novel. We have had our share of people who, in spite of all disclaimers, seem really to have expected that marvelous results might follow easily once their data were somehow "computerized." Naturally nothing really marvelous has come out of these early studies. Results have always been interesting, and in some cases important contributions to archaeological problems have been made. Yet, nothing done so far has convinced the archeological profession as a whole that there are any often-encountered tasks or problems for which computers ought to be used as a matter of course; that there are tasks for which it would show incompetence *not* to use a computer. My impression is that the majority of archaeologists are still watching results of computer work with attitudes ranging from hostility to friendly interest, and are not going to make any real commitment either to learning or to using machine techniques until there is more evidence that computers can really offer economies in the performance of familiar tasks, or that the results of novel computer approaches are really valid and intelligible. The incidence of intelligent comprehension of computers is still depressingly low among archaeologists of all ages, although it may be rising rapidly in the current generation of students. I think archaeologists really engaged in computer work have reached a "second generation" stage where it is more fully appreciated that a great deal of hard work, hard thinking, and trial and error are still needed before we can make the best uses (and non-uses) of computers, but nonetheless a stage where we have a substantial body of earlier efforts whose successes and failures we can learn from. It seems wasteful for either archaeologists or computer people trying to help archaeologists to begin computer projects today without knowing what has already been done or attempted in archaeology. An extremely important source of information on this work is the *Newsletter of Computer Archaeology*.

Important tasks

Most archaeological applications of computers fall

*The present paper has profited from comments on an earlier version made by R. A. Benfer, Junius Bird, R. G. Chenhall, Keith Dixon, J. C. Gardin, Dee F. Green, James Hill, F. R. Hodson, T. Hunter, Jesse D. Jennings, William Lipe, I. Scollar, A. C. Spaulding, S. Sedelow, and Bruce Warren. They have informed me of some new work, made some stylistic suggestions, and indicated places where my meaning was unclear. I gratefully acknowledge their help.

under the broad headings of data storage and retrieval or of multivariate statistical analyses. The two most important tasks we face with respect to these applications are developing optimum procedures for coding our data, and getting a better control of mathematics. Both involve much that can be done only by the archaeologists themselves; for only the people working with the data can say what kinds of things may be important to record about it, and certainly no one else can learn our math for us. This last point needs some stressing because archaeologists rarely know anything beyond high school algebra, and most do not know even that really well. My impression is that most archaeologists without computer experience believe that any competent programmer is also an all-round mathematical authority and do not realize that they may need important statistical advice which a programmer cannot give them. Probably programmers for their part often assume that archaeologists know more than they in fact do about the best ways to formulate or to solve their problems. Also, archaeologists who have gotten some kind of statistical answer out of a machine have been at best shaky about what to make of it once they had it, and in some cases have been confused or misled by the results. Whatever other roles computers come to play in archaeology, we will not be able to get very much out of statistical methods until there is a higher level of mathematical competence among archaeologists. Minimal requirements are that archaeologists be able to judge whether a suggested technique is really appropriate for their data and problems, and that they recognize the need for competent statistical advice as something distinct from programming assistance. It is also desirable that some one develop new, more suitable mathematical techniques in cases where nothing now existing is really quite what we want.

Archaeological data codes

The kinds of codes we especially need to develop are those for describing physical objects, both for data storage and retrieval and for statistical analyses. This is particularly so for broad classes of materials such as pottery, chipped stone, ground or polished stone, metals, and fibers and textiles. It can also be useful to encode patterned concatenations of objects, for example structures or burial arrangements. For any specific study, the most efficient coding will depend on the range of variations shown in a specific body of data. But for each of the broad classes of materials mentioned above, there are substantial universal similarities in what it is relevant to describe and in the limits of variation. It is certainly undesirable that every computer project develop a completely *ad hoc* coding scheme. This is not to say that we should not continue to explore and experiment with new approaches to old coding problems, nor that special codes will not be desirable for special data or special problems. Above all I do not mean that we should commit ourselves to some kind of "least common denominator" framework which will lend itself mainly to coding features which occur widely in many different cultural traditions. On the contrary, a major advantage to computer techniques should be a greatly improved ability to deal with fine distinctions in detail relevant only for the quite specific and localized entities which Lipe[5] calls "micro-traditions." The point is that there are kinds of features of near-universal importance which ought to be dealt with well by our codes, whatever else in addition they handle. Also, there are recurrent problems in expressing the way in which parts of an object relate to one another to form the whole object (what, to use the linguistic kind of approach favored by Gardin, might be called "syntactic" relationships). People starting new computer projects will very likely find that many of the problems they face have had to be faced in earlier projects, and they ought to consider earlier attempted solutions carefully and adopt a consistent scheme unless they can think of something to try which offers a possibility of working better. Clearly we should continue to try out many new things just to see how they will work, but we do not need to re-solve already solved problems or to duplicate old mistakes. General frameworks and standardized practices will save wasted effort and will help toward making the best balances between particularity and universality in specific codes; without requiring that codes for every data file for a given kind of material be highly similar. Important recent general discussions of coding include those by Gardin[1] and by Chenhall.[6]

Archaeologists are currently debating among themselves about the degree to which the features we find important for establishing categories of artifacts, and the categories themselves, could ever be demonstrated to be close to those really "in the minds" of the dead makers and users of those objects. Whether or not this will ever be possible, at least we should not put unnecessary obstacles in the path toward such a goal. Furthermore, for a wide variety of statistical techniques, we need to be able to generate numbers from the basic data file which will be reasonable expressions of the resemblance to one another of objects or of sets of objects; based either on all their recorded features or on some definite subset of their features. Both of these considerations imply that descriptive codes should always be reasonably related to human judgments made by workers experienced with the corpus of objects. Features judged to be quite similar should have this similarity reflected in their coding, and similarities in "syntactic" arrangements of features on objects should also be reflected, even though the specific features may differ. Emphatically, this does not mean that we should merely find ways of translating currently popular descriptions into a code readable by machine, still less does it mean we should arrive at the same categories as before. Current archaeological description practice is to omit much detail because it is simply unmanageable, and very often categories are quite palpably derived by selective emphasis on certain features at the expense of others, or by methods which frankly strike other workers as idiosyncratic. I also do not mean to say that logical clarity guarantees good results, for it is easy to do something which is logically rigorous and consistent and absurd. We need a fine interplay between human insight and common sense and the vast literal-minded idiotic power of the machines. All this may seem to be stating the obvious, but I mention it, especially the need that the original data file use a code related to our perceptions about similarities in features and similarities in configurations of features, because it is quite possible to use codes which lack these properties. For example, one could easily store outlines of objects by listing a series of grid coordinates, but such a system makes it hard to emphasize differences at certain key points of the profile, or to express the strong fundamental similarities in objects which differ only moderately in some proportions.

There may be important lessons for us in the work being done on pattern-recognition by machine, but so far I know of no useful archaeological applications of these techniques. To describe the geometrical aspects of solid objects like pottery, stone, or metal, what I think we need are some important size measurements, locations of often-relevant "landmark" points, information about profiles at and between these points, symmetry properties, and information about location and character of special attachments to or deletions from the main body. Important non-geometrical kinds of information include raw materials, clues about methods of manufacture, clues about use, features of finish and decoration, context in which the object was found, and catalog number or other positive identification of the object. Chenhall has discussed these matters in an unpublished M. A. thesis,[7] although unfortunately his published work[6] does not treat coding shapes except in very general categories.

For geometric description of pottery vessels, I think the broad outlines of a good approach are clear. This is most fully described by A. O. Shepard.[8] McGimsey and Green,[9] and Gardin[1] illustrate very similar approaches. What is needed here is not some drastic revision, but improvement and standardization of this general method, insuring that it retains an "open-ended" quality which will permit its application to the details of any conceivable ceramic tradition. I am less in touch with work being done on chipped stone tools, but I believe that there is fair agreement about what kinds of things it is important to notice and less agreement on a standard way to encode these observations. Important work here includes the lithic typology conference reported by Krieger[10] and Weyer,[11] and the codes described by Binford[12] and Jennings.[13] Stern[14] gives a code for ground stone. Christophe and Deshayes[15] present a code for metal tools and weapons. An unpublished code for computer analysis of textiles, developed at the American Museum of Natural History, is discussed by Bird.[16]

The matter of developing general codes for decorations of objects, where the concern is with design elements, style, iconographic content, or subject matter, seems far more difficult. It is an area where methods adapted from descriptive linguistics seem very promising. Gardin[1] gives a very important discussion and illustration of some of these methods. Also very important, though not intended for immediate use with computers, is the work of John Rowe and his students at Berkeley,[17] and of Muller.[18] This is plainly a topic where archaeology and art history have many similar needs and problems.

One application of archaeological data codes which is somewhat distinct from their use in specific research projects is in the "computerization" of the catalogs of large museums. This task is under way or seriously projected by Dee Green at the University of Missouri and by Jaime Litvak and Felicity Thomas at the Instituto Nacional de Antropologia e Historia in Mexico City. Irwin Scollar [19] reports that this may be done at the Rheinisches Landesmuseum in Bonn.

Statistical studies

One reason why computers have not yet had any great impact on archaeological practice is that no one has yet completed and made available a file containing any really large body of important data coded to include what is relevant for important problems. Probably the largest published file is by Christophe and Deshayes,[15] which includes about 4000 metal objects, using optical coincidence cards rather than electronic equipment. The statistical studies which have been done so far have never been based on samples of more than a few thousand objects (at most, a few tens of thousands if very fragmentary objects or workshop debris are included). In many cases, of course, good samples of this order of magnitude are quite sufficient to produce important results, but it does mean that the volume of data processed has never been very large relative to the millions of objects (mostly small pottery fragments) which major excavations often produce. The importance of computer studies will increase greatly when larger files of significant data are accumulated, especially as whole regions and substantial time spans come to be covered with some adequacy. Even so, it does not seem profitable to try to include all data on all objects excavated, and good statistical sampling design is a matter of increasing concern.[20]

Some important examples of archaeological applications of statistical techniques by computer include theuse of chi-square and regression by Freeman and Brown;[21] multiple regression by McPherron[22] and by Longacre;[23] factor analysis by Jennings,[24] Binford and Binford,[25] Cowgill,[26] Hill,[27] and Benfer,[28] proximity analysis by Hodson and others;[29] scalogram methods by Eliseeff;[30] and automatic classification methods by Hodson[29] and De La Vega.[31]

Few archaeologists have been involved in development of new statistical programs for their work. Deetz's[3] work in 1960 is one exception. For a number of reasons, his work was probably more influential than any other single project in persuading American archaeologists that computers might possibly be valuable for them. Deetz addressed himself to an original and important problem; whether the historically documented breakdown of social organization (particularly a pattern of matritocal residence) under increasing European pressures on an Indian village in South Dakota in the 18th century might be reflected in a parallel breakdown of clustering in ceramic design elements. But it is evident that no one with basic statistical competence gave this work any serious attention. Many archaeologists, notably Deetz himself, are well aware that even well-demonstrated changes in clustering of artifact design elements may not have clearcut social implications. What needs to be emphasized is that Deetz's demonstration method itself involved computations that were unnecessarily tedious and unnecessarily ambiguous, and should not be used as a model for further work. He attempted to assess degrees of association among cross-tabulated attributes, by a technique which was ingenious but less useful than standard measures like phi or lambda.[32]

Kuzara, Mead, and Dixon[33] have developed what seems to be a very good program for the task known to archaeologists as "seriation"—arranging a set of units in the order which best satisfies the requirement that the more similar any two units are, the closer to one another they are in the final sequence. Archaeologists have had a fair amount of experience in doing this directly by manual rearrangement of the units, so there is already fair understanding of the rationale, and the convenience of doing it by computer is appreciated. Kuzara, Mead, and Dixon's program appears to work better than an earlier one designed for this purpose by the Aschers.[34] It has already been applied by other archaeologists, including a study of stone tools in Texas by LeRoy Johnson,[35] and it has good prospects of becoming quite popular. Its only serious limitation is that it amounts to ordering units along some one best axis or factor. Where there is any reason to think two or more factors may be relevant, it would be preferable to use a multidimensional technique such as factor analysis or something like R. Shepard's proximity analysis.[36]

In most multivariate approaches a pervasive theme is the drive toward certain kinds of parsimony. What are the best variables for discriminating between members of several categories, what are the best predictor variables for some set of criterion variables, or what are the fewest independent factors which account for most non-random variance in some larger set of variables? In many archaeological problems these are indeed the kinds of parsimony we want. Often, though, we really want something else; namely, the most parsimonious account of the patterning of *all* variables of some set. Probably we should rely less on methods developed for the reduction of experimental data (especially by psychologists, agronomists, and biologists), and more on analogies with descriptive grammars. Lounsbury's[37] approach to formal accounts of systems of kinship terminology is an especially important and lucid exposition of this approach. Excellent archaeological work along these lines has been done by John Rowe and those influenced by him at Berkeley,[17] working without computers. Rowe and his students have produced results far more important than anything which has been done so far in archaeology by computer, largely because they have applied a good method to rich bodies of data, while computer studies have been short on one or both of these scores. Muller[18] has also done important work in applying a generational grammar approach to a prehistoric art style. It is likely that computers could be used to make the "grammatical" approach more powerful and less laborious, but this will require more hard and original thinking, than is demanded by the adoption of ready-made multivariate programs. Gardin's[1] work on codes for iconography is an important contribution in this direction. Sackett's[38] non-computer work on multiple contingency tables is superficially quite different, but is probably also leading in the same direction.

The general field of mathematical geography, or mathematical analysis of spatially distributed data, is another promising field for computer applications to archaeological data. Work here includes Lipe and Huntington's use of centrographic techniques for demonstrating differences in distribution of ceramic categories,[39] and the use of linear spatial filtering to improve contrast in plots of magnetometer survey data by Scollar and Krückeberg.[40] Data smoothing and trend surface fitting techniques are also likely to prove useful. I am currently engaged in analysis of data from Teotihuacan, a 25 square kilometer prehistoric metropolis in central Mexico,[41] where differences between districts within the city is one major concern. We are working with a data matrix of 391 observations on each of possibly 4000 units. While most of our computer work utilizes multivariate statistical methods, we have also found it useful to produce maps by computer of data distributions using a program (SYMAP) developed under the direction of Howard T. Fisher, of the Laboratory for Computer Graphics of Harvard University. By far the greatest advantage over hand methods comes when functions of data at two or more points must be computed, as in smoothing or filtering procedures.

Other applications

One special field of computer work is on decipherment of ancient writing systems. An early attempt to decipher Maya hieroglyphs by Evreinev, Kosarev, and Ustinov at Novosibirsk was unsuccessful and strongly criticized by others, including Knorozov.[42] Current work in Mexico on a concordance of Mayan inscriptions is not aimed toward instant decipherment and is likely to be far more useful.[43] At least two computer projects involving Minoan writing are presently under way.[44]

A KWIC index of *American Antiquity,* a major American journal, has been produced by Dee F. Green[45] but is not yet published. According to Irwin Scollar[19] the annual and cumulative indexes of the *Bonner Jahrbuch* and a concordance of aerial photos of archaeological sites are all being compiled by computer at the Rheinisches Landesmuseum, Bonn.

Perhaps the most unusual computer application so far in archaeology is in connection with the work of G. Hawkins, who used computed ancient stellar positions for his study of the astronomical significance of Stonehenge. James Dow has also used this program for research on possible stellar bases for orientations of ancient cities and temples in Mexico.[46] Undoubtedly many more special applications of computers in archaeology will appear, in addition to their major uses for data storage and retrieval and for statistical and formal analysis of data.

REFERENCES

1 J C GARDIN
Methods for the descriptive analysis of archaeological material
American Antiquity 32 13-30 1967

2 J C GARDIN
Reconstructing an economic network in the ancient East with the aid of a computer
In Hymes D The Use of Computers in Anthropology
Mouton & Co The Hague pp 377-391 1965

3 J DEETZ
The dynamics of stylistic change in Arikara ceramics
University of Illinois Press Urbana 1965

4 R G CHENHALL (editor)
Newsletter of computer archaeology
Department of Anthropology Arizona State University
Tempe 1965 and later
Other useful sources include:
D HYMES
The use of computers in anthropology
Mouton & Co The Hague 1965
B WARREN
Computers and research in archaeology
Dittoed 1965
D F GREEN
Computer bibliography
Machine listing available on request from the author
Weber State College Ogden Utah 1967

5 W D LIPE
Personal communication
1967

6 R G CHENHALL
The description of archaeological data in computer language
American Antiquity 32 161-67 1967

7 R G CHENHALL
An investigation of taxonomic systems for the storage and retrieval of material-culture data on electronic computers
Department of Anthropology Arizona State University
Tempe 1965

8 A O SHEPARD
Ceramics for the archaeologist
Carnegie Institution of Washington Washington DC
1957

9 C R McGIMSEY D F GREEN
IBM ceramic code outline
University of Arkansas Museum 1965

10 A D KRIEGER
New world lithic typology project: part II
American Antiquity 29 489-493 1964

11 E M WEYER
New world lithic typology project: part I
American Antiquity 29 487-489 1964

12 L R BINFORD
A proposed attribute list for the description and classification of projectile points
University of Michigan Anthropological Papers 19 193-221 1963

13 J D JENNINGS
Information on University of Utah computer study
Department of Anthropology University of Utah 1964

14 E M STERN
Using the IBM 7090 in the classification of ground stone tools
Michigan Archaeologist 12 229-234 1966

15 J CHRISTOPHE J DESHAYES
Index de l'outillage sur cartes perforées: outils de l'age du bronze, des Balkans á l'Indus

Centre National de la Recherche Scientifique Paris 1964

16 J BIRD
The use of computers in the analysis of textile data; specifically archaeological fabrics from Peru
The American Museum of Natural History New York 1967

17 R P ROARK
From monumental to proliferous in Nasca pottery
Ñawpa Pacha 3 1-92 Dept. of Anthropology University of California Berkeley 1965

18 J MULLER
Style and archaeology
Department of Anthropology Southern Illinois University Carbondale 1967

19 I SCOLLAR
Personal communication 1967

20 Important recent papers on archaeological sampling include Vescelius G *Archaeological sampling a problem of statistical inference*
in Dole and Carneiro Essays in the Science of Culture in Honor of Leslie A White Thomas Y Crowell New York 1960 pp 457-70
S ROOTENBERG
Archaeological field sampling
American Antiquity 30 181-188 1964
G L COWGILL
The selection of samples from large sherd collections
American Antiquity 29 467-474 1964
L R BINFORD
A consideration of archaeological research design
American Antiquity 29 425-441 1964
J N HILL
Random sampling a tool for discovery
Department of Anthropology UCLA 1967
clude Vescelius G

21 J A BROWN L G FREEMAN
A UNIVAC analysis of sherd frequencies from the Carter Ranch Pueblo Eastern Arizona
American Antiquity 30 162-167 1964
L G FREEMAN J A BROWN
Statistical analysis of Carter Ranch pottery
Fieldiana: Anthropology 55 126-154 Chicago Natural History Museum 1964

22 A McPHERRON
Programming the IBM 7090 for optimizing taxonomy in archaeology
Department of Anthropology University of Pittsburgh 1963

23 W A LONGACRE
Archaeology as anthropology: a case study
Science 144 1454-55 1964

24 J D JENNINGS *op cit*

25 L R BINFORD S R BINFORD
A preliminary analysis of functional variability in the Mousterian of Levallois facies
American Anthropologist 68 no 2 part 2 238-295 1966

26 G L COWGILL
Evaluación preliminar de la aplicación de métodos a máquinas computadoras a los datos del mapa de Teotihuacán
Department of Anthropology Brandeis University Waltham Mass 1966

27 J N HILL
A prehistoric community in eastern Arizona
Southwestern Journal of Anthropology 22 9-30 Albuquerque 1966

28 R A BENFER
A design for the study of archaeological characteristics by population genetical and psychological models
American Anthropologist in press

29 F R HODSON P H A SNEATH J E DORAN
Some experiments in the numerical analysis of archaeological data
Biometrike 53 311-24 1966
J E DORAN F R HODSON
A digital computer analysis of Palaeolithic flint assemblages
Nature 210 688-89 1966

30 V ELISEEFF
Possibilités du scalogramme dans l'étude des bronzes chinois archaiques
Mathematiques et Sciences Humaines 11 1-10 1965

31 W F DE LA VEGA
Classification des tombes d'une nécropole d'Italie du sud, sur calculateur
Proceedings of the International Symposium on Computational and Mathematical Methods in the Behavioral Sciences Rome 1966 International Computation Centre Rome in press

32 Much the same point is made in a review of Deetz's work by Spaulding A C *American Anthropologist* 68 1064-5 1966

33 R S KUZARA G R MEAD K A DIXON
Seriation of anthropological data: a computer program for matrix-ordering
American Anthropologist 68 1442-55 1966

34 M ASCHER R ASCHER
Chronological ordering by computer
American Anthropologist 65 1045-52 1963

35 L JOHNSON
Towards a statistical overview of the archa cultures of Central and Southwestern Texas
Texas Memorial Museum Austin Bulletin 12 1967

36 R N SHEPARD
The analysis of proximities: multidimensional scaling with an unknown distance function
Psychometrika 27 125-140 and 219-246 1962

37 F G LOUNSBURY
A formal account of the Crow- and Omaha-type kinship terminologies
In Goodenough W editor Explorations in Cultural American Anthropologist 68 no 2 part 2 356-94 York 1964

38 J R SACKETT
Quantitative analysis of Upper Palaeolithic stone tools
Department of Anthropology State University of New 1966

39 W D LIPE C F HUNTINGTON
The application of some centrographic techniques to the analysis of archaeological data
Anthropology McGraw-Hill Book Company Inc New York at Binghamton 1964

40 I SCOLLAR F KRÜCKEBERG
Computer treatment of magnetic measurements from archaeological sites
Archaeometry 9 61-71 1966

41 Principal investigator of this project is René Millon
Papers concerning the computer work include
G L COWGILL
Computers and prehistoric archaeology
In E Bowles editor Computers in Humanistic Research Prentice-Hall Inc Englewood Cliffs New Jersey 1967 chap 6 pp 47-56
G L COWGILL
Computer archaeology at Teotihuacan Mexico 1965
Statistical and computer approaches to sociocultural interpretation of an ancient city of Mexico 1966
Evaluación preliminar de la aplicación de mátodos a máquinas computadoras a los datos del mapa de Teotihuacán
Department of Anthropology Brandeis University Waltham Massachusetts 1966

42 IU V KNOROZOV
Machine decipherment of Maya script
Soviet Anthropology and Archaeology 1 43-50 1962/3

43 J J RENDÓN A SPESCHA
Nueva clasificación plastica de los glifos Mayas
Estudios de Cultura Maya 5 189-280 Mexico City 1965

44 One is by Elizabeth W Barber Department of Linguistics Yale University The other is by Richard Morgan and John Reich Classics Dept University of Manitoba cited in Computers and the Humanities 1 233 1967

45 D F GREEN personal communication 1967

46 J W DOW
Astronomical orientations at Tectihuacán a case study in astro-archaeology
American Antiquity 32,326-334 1967