

VOLUME 1 NR. 2 | AUGUST 1968

Reprint

COMPUTER
STUDIES

*in the Humanities
and Verbal Behavior*

MOUTON

COMPUTER STUDIES

in the Humanities and Verbal Behavior

CO-EDITORS

FLOYD R. HOROWITZ, University of Kansas
LEWIS SAWIN, University of Colorado
SALLY Y. SEDELOW, University of North Carolina

BOARD OF EDITORS

<i>Anthropology</i>	GEORGE L. COWGILL, Brandeis University	<i>Mathematics</i>	PAUL R. HALMOS, University of Michigan
<i>Archaeology</i>	PAUL S. MARTIN, Field Museum of Natural History	<i>Music</i>	LEJAREN HILLER, University of Illinois
<i>Art</i>	LESLIE MEZEI, University of Toronto	<i>Philosophy</i>	LARRY TRAVIS, University of Wisconsin
<i>Bibliography</i>	ERIC BOEHM, American Bibliographical Center, Santa Barbara	<i>Political Science</i>	OLE R. HOLSTI, University of British Columbia
<i>Classics</i>	T. M. ROBINSON, University of Calgary	<i>Programming</i>	DAVID BRIDGER, Washington University
<i>Education</i>	ROBERT SCHMIEL, Goucher College		DANIEL E. BAILEY, University of Colorado
<i>Folklore</i>	ELLIS B. PAGE, University of Connecticut	<i>Psycholinguistics</i>	JOHN B. CARROLL, Educational Testing Service, Princeton
<i>History</i>	JOHN Q. ANDERSON, University of Houston	<i>Social Psychology</i>	HANS E. LEE, Michigan State University
<i>Library Science</i>	THEODORE K. RABB, Princeton University	<i>Sociology</i>	WALTER A. SEDELOW, University of North Carolina
<i>Linguistics</i>	RALPH H. PARKER, University of Missouri		EDWARD E. DAVID, Bell Telephone Laboratories
<i>Literature and Modern Languages</i>	SHELDON KLEIN, University of Wisconsin	<i>Speech</i>	GERALD M. SIEGEL, University of Minnesota
	ROBERT S. WACHAL, University Of Iowa		JULIET SHAFFER, University of Kansas
	BERTRAND AUGST, University of California	<i>Statistics</i>	HAROLD P. EDMUNDSON, University of Maryland
	JESS BESSINGER, JR., New York University		GARY GAISER, Indiana University
	WILLIAM INGRAM, University of Michigan	<i>Theater Translation</i>	SILVIO CECCATO, Università di Milano
	HENRY KUERA, Brown University		DAVID A. DINEEN, University of Kansas
	JAMES W. MARCHAND, Cornell University		
	STEPHEN PARRISH, Cornell University		
	ALICE POLLIN, New York University		
<i>Mass Communication</i>	WILLIAM J. PAISLEY, Stanford University		

EDITORIAL ADDRESS

Computer Studies in the Humanities and Verbal Behavior is jointly sponsored by the universities of Colorado, Kansas and North Carolina. Manuscripts: LEWIS SAWIN, Department of English, University of Colorado, Boulder, Colorado, U.S.A. General correspondence: FLOYD R. HOROWITZ, Department of English, University of Kansas, Lawrence, Kansas, U.S.A.

Computer Studies in the Humanities and Verbal Behavior appears quarterly in issues of approximately sixty pages. Four issues constitute a volume. The subscription price is \$ 10.00/f 36.00 per year per volume; single issues cost \$ 3.00/f 10.50. Back issues as well as subscriptions and single issues can be ordered from every bookseller or subscription agency, or directly from Mouton & Co., P.O. Box 1132, The Hague, The Netherlands.

The Computer in Archaeology

F. T. CLOAK, JR.

University of North Carolina

The author's primary thesis is that the computer can be used more effectively in archaeological studies than has heretofore been the case. He first presents a capsule review of some kinds of work already done by computer, including a description of Albert C. Spaulding's research ("Statistical Description and Comparison of Artifact Assemblages"). He next considers the usefulness of the computer in analyzing data obtained by means of stratigraphy and seriation. As a substitute for traditional seriation methodology the author then suggests ways in which the computer can be used to determine whether a given type of attribute occurs in an assemblage and to distinguish developmental, orderly change from random or accidental change. This suggestion is based on the author's research on a Trinidadian village, the methodology of which is described in detail. Ultimately, the author believes, the computer may help the archaeologist move from primary questions ("What was in the producer's mind?" and "How did it get there?") to still another level of humanistic concern, "By what processes did the culture that put it there come to have it?"

The use of the high-speed computer in archaeology is still in its infancy, one might even say in its gestation period. So far, computers have been used merely to apply statistical techniques that were worked out for use with pencil and paper and desk calculators. A friend of mine in mathematical statistics tells me that the trend in that field is away from the development of indices and coefficients of more-or-less general applicability, and toward the use of what are sometimes called "empirical procedures", making more direct use of the peculiar capabilities of the computer to solve the specific problem at hand. I refer to such operations as Monte Carlo, hill climbing, and iterative procedures in general.'

In other words, we are not getting the most out of computers in archaeology. To give an example: Another friend of mine, working on a problem in archaeology, wanted to classify many thousands of potsherds he had gathered in Mexico. This required that a certain coefficient be calculated for 40 different groups of sherds. Working with a desk calculator, he calculated one of these coefficients. It took him one week of arduous labor; thus the 40 operations would have taken him 40 weeks. He turned to a computer, which did all 40 in 15 seconds.

* Read at the Conference on The Computer and Research in the Humanities, at the University of North Carolina, March 29, 1967. ¹ For a general discussion of these procedures, see Bert F. Green, Jr., *Digital Computers in Research: An Introduction for Behavioral and Social Scientists* (New York, McGraw-Hill, 1963), pp. 149-158.

I had a problem with a similar amount of data, in this case ethnographic data I collected in Trinidad. For my problem there were no statistical procedures available. It is unlikely, in fact, that routine statistical procedures will ever be developed for this problem. I turned to the same computer, using the same software, utilizing some empirical procedures which I will outline for you later. The total running time for my project was 25 MINUTES, one hundred times as long as my friend's. On that basis, it would have taken 4,000 weeks, or about 80 years, to do the job by hand. At the same time, natural science projects run at the same computing facility were taking several hours — on a much faster machine. The point is, my friend's project could have been done in pre-computer days. Mine could have been done, but it wouldn't have because of the tremendous cost in labor. The natural science projects, no matter how important, simply couldn't have been done at all.

Actually, what I have said above does not tell the whole story. My friend, by employing 40 clerks with 40 desk calculators, could have gotten his job done in one week. My job, on the other hand, was such that each step required the use of the results of the previous step so, while each step might have been farmed out to several hundred clerks, the total time would have had to be a matter of some months at best. The difference is really one of kind, not merely of degree. And I dare say that the difference

between my job and those being run by the physics department was one of kind also. We haven't really begun to use the computer, as a computer, at all.

The discussion of my project is continued below, after a review of the kinds of archaeological problems computers have been used for.

Modern archaeology is frequently referred to by its practitioners as "scientific" archaeology, because of the sophisticated and meticulous methods which are used in gathering and processing data. I would not dispute the appellation, but would only add that modern scientific archaeology has in no sense removed itself from the humanities. The primary purpose of any archaeological study is to examine the product of a man's or woman's hands in order to determine what was going on in the MIND of that man or woman when that product was produced. The next purpose is to determine how that "what-was-going-on" came to BE in the mind of that man or woman. These two questions — What was in the producer's mind? How did it get there? — are the very questions asked by literary critics, art historians, and all other students of the humanities.

Let us examine, then, how modern scientific archaeology approaches these two questions, with an eye to computer applications.

Much of the following is based on an article entitled "Statistical Description and Comparison of Artifact Assemblages", by Albert C. Spaulding.² Spaulding discusses the raw materials of archaeological analysis, as follows:

Archaeologists typically direct their attention to sets or groups of artifacts rather than to single objects, a notable contrast to the attitude of students of the humanities. [Let us say, "The *rest* of the humanities".] The foundation for this interest in collections is essentially taken for granted here and [is] implied without explanation by the use. . . of such terms as "cultural behavior" and "culturally relevant". Culture is patterned behavior at the level of symbolic activity, and the patterning and symbolism are interpersonal, a product of human social life. Artifacts tend very strongly to occur in the spatial clusters we call sites primarily because their makers and users lived in societies. The ideal unit of archaeological study is the assemblage of artifacts produced and used by a single society over a period of time short enough to preclude any marked changes through cultural innovations or shifts . . . ³

² Albert C. Spaulding, "Statistical Description and Comparison of Artifact Assemblages" in Robert F. Heizer and Sherburne F. Cook (eds.), *The Application of Quantitative Methods in Archaeology* (= *Viking Fund Publications in Anthropology*, No. 28) (Chicago, Quadrangle Books, 1960), pp. 60-83. Copyright 1960 by the Wenner-Gren Foundation for Anthropological Research, Inc. Used by permission of Quadrangle Books, Inc.

³ *Ibid.*, p. 61.

Spaulding then proceeds to what is logically the first step in the analysis of an archaeological collection — the identification and definition of the culturally relevant ATTRIBUTES (qualities, properties) of the various artifacts. As an example, he uses the determination of the frequency distribution of the lengths of the projectile points from a certain assemblage or collection. This distribution of lengths is shown to be very close to a NORMAL distribution, a

systematic pattern of the single-peaked type, a pattern which is common in lists of measurements of the most diverse phenomena. This orderliness suggests strongly that one powerful centralizing tendency is at work but that each individual expression of the tendency is more or less modified by a number of minor factors operating in a random fashion. I would infer that the centralizing force was a culturally imposed ideal length for our class of projectile points and that this ideal length is estimated by the mean of the distribution. Variations from this mean presumably resulted from such factors as individual discrepancies in visualizing the ideal length, accidents and imprecision in the stone chipping process, and lack of homogeneity in the raw material. The combined strength of these modifying factors is estimated by the standard deviation. From the standpoint of cultural patterning, we have clear evidence that only one metrical attribute is represented; all of the projectile points of the particular class under consideration were *intended* to be about the same length, and there is no possibility of subdividing the class meaningfully into, for example, long and short categories. If the curve fitting had shown that the distribution was not normal, it would be necessary to reexamine the data for possible explanations. A non-normal curve might show more than one peak, with the resulting inference that there were two or more ideal-lengths, possibly implying functional differences, or it might appear to be asymmetrical, suggesting that the intended use imposed a fairly sharp restriction on how long or how short the point could be, and so on.⁴

The determination of mean, standard deviation, and the probability that the distribution was normal is, of course, a very simple operation on a computer.

In another example, Spaulding shows a simple statistical method for determining whether the use of a certain tempering material is deliberate or fortuitous, by testing its association with another attribute, the use of a red surface finish, which is obviously intentional.⁵ Again, the

⁴ *Ibid.*, p. 66. It should be noted here that the result of these procedures is to convert QUANTITATIVE data into QUALITATIVE data; the attribute, that is, is no longer "metrical" but "discrete", to use Spaulding's terms. A particular point can now be described as exhibiting or as not exhibiting a particular feature of length. I believe that all cultural elements are essentially of the presence-or-absence type, as are biological elements; a given "unit of cultural instruction" is either transmitted or not transmitted from one individual to another. Cf. F. T. Cloak, Jr., "Cultural Microevolution", *Research Previews* (Institute for Research in Social Science, University of North Carolina), Vol. 13, No. 2 (November, 1966), pp. 7-10.

⁵ Spaulding, *loc. cit.*, pp. 67-68.

objective is to find out what the artisan had in mind — Did he intend to use limestone temper or did it just happen?⁶

When the identification and definition of the culturally relevant attributes is complete, the next step, logically, is the classification of the artifacts on the basis of these attributes. Sometimes classes or types are perceived intuitively by the archaeologist; when this is the case, the validity of these types can be checked by association tests of various kinds. A computer may be useful for carrying on the computations involved.'

At other times, however, the amount of material and the number of possible types may be so great that the archaeologist's intuition fails. In this case, so-called "automatic classification" procedures such as cluster analysis and factor analysis may be called for. Such procedures would really make use of the special capabilities of the high-speed computer but as recently as 1965, no one had yet written a program for automatic classification suitable for the qualitative type of data characteristic of archaeology.⁸

Whatever techniques are used for the development of classes or types of artifacts, however, the resulting classification presumably reflects the mental set of the artisans who were producing those artifacts. As Spaulding puts it :

My principal point is that the symbols of the analysis stand for attributes judged to be culturally meaningful on the basis of all the wisdom and experience that the archaeologist can bring to bear on the problem, and it is a fair inference that significant relationships between culturally meaningful entities have cultural and hence behavioral significance. If an expanding stem on the base of a flint projectile point is a good attribute, and if small size of flint projectile points is a good attribute, and if small size and an expanding stem are closely associated in the collection of projectile points, then our best explanation for the association is that the makers reproduced a customary pattern — that they thought of the expanding stemmed and small projectile point as a definite sort of projectile point. Further, we are permitted to suspect that the special kind of projectile had a special function or range of functions and perhaps even that the makers had a special name for the type.⁹

⁶ For a list of proposed universal, intersubjective attributes, see Jean-Claude Gardin, "Four Codes for the Description of Artifacts: An Essay in Archeological Technique and Theory", *American Anthropologist*, Vol. 60 (1958), pp. 335.

⁷ For a brief history of quantitative type-coding methods, with bibliography, see Harold E. Driver, "Survey of Numerical Classification in Anthropology", in Dell Hymes (ed.), *The Use of Computers in Anthropology* (The Hague, Mouton, 1965), pp. 320-323.

⁸ Wilhelm Milke "Statistical Processing", in Hymes, *loc. cit.*, p. 195. See also Peter Ihm, "Automatic Classification in Anthropology" in Hymes, *loc. cit.*, pp. 357-376.

⁹ Spaulding, *op. cit.*, p. 76.

Before we leave the discussion of artifact classification by attributes, we should mention that there is a definite analytical feedback between this classification level and the lower level of attribute determination. That is to say, our attempts to classify attributes together into types may make us realize that what we had thought was one attribute was in fact two, or that what we had thought were two attributes was really a single attribute with a rather large amount of free variations within it, or even that what we had thought was a cultural attribute was in fact controlled by non-cultural factors. (For instance, we might find that the limestone temper in the example mentioned above was actually an accidental inclusion after all.)

We now turn to the classification not of artifacts but of assemblages of artifacts. Assemblages may be classified by the types of artifacts found in them, by the attributes of those artifacts directly, or by a combination of both types and attributes."

The technical difference between classification of artifacts and classification of assemblages is that while an artifact either has or lacks a given attribute, an assemblage has a certain FREQUENCY of a given type of artifact (or of a given attribute). This frequency is generally expressed as a percentage of the total number of types or attributes, and an index of likeness is calculated for each pair of assemblages ; in other words, each assemblage is compared with every other assemblage, and their similarity is expressed numerically. If the degree of similarity is found to be improbably high, the assemblages are deemed to be of the same culture. Thus a large number of assemblages can be sorted into cultures, implying that the people who produced those assemblages were influenced by the same, or different, or more or less closely related, cultural traditions.

Again, the number of computations involved makes a computer extremely useful, although only a fraction of the potential capacity of the computer would be used.

We have now moved to the second of the humanistic goals of archaeology mentioned before. The first was ascertaining what was in the mind of the producer; the second is ascertaining how it got there, satisfied here by specifying the cultural tradition of which the producer partook. At this point we could say, hopefully, that a certain group of flint-knappers intended to make their projectile points about 9.3 centimeters long and put such-and-such a style of tang on them, and so on, because they were bearers of culture C 1; had they been bearers of culture C 2, they would have tried to make them 7.1 centimeters long.

Cf. Irving Rouse, *Prehistory in Haiti: A Study in Method* (= Yale University Publications in Anthropology, No. 21) (New Haven, Yale University Press, 1939).

meters long and to put a different style of tang on them.

But the notion of a cultural TRADITION implies a historical development, of which the culture of our flint-knappers is merely the end-product at a particular moment in time. Archaeologists are particularly interested in studying the development of cultural traditions. This is done, of course, by comparing assemblages which are the end-products, or residues, of a SINGLE cultural tradition at VARIOUS moments in time. The ideal method of studying cultural change through time is, of course, through stratigraphy. If a people occupied the same site for an extended period of time, say hundreds of years, and if conditions of deposition were just right, we can come along much later and find their assemblages of artifacts neatly stacked one on top of another, with the oldest in the bottom. We can then analyze each layer and determine when a new type or attribute was first made and how its popularity waxed and waned over the years, and whether it was eventually given up altogether, and thus trace the cultural history of the people.

Suppose, however, the people moved around a lot, staying at each site only a year or two or ten. Although we might have dozens of assemblages, we would have no stratigraphy, no direct way of determining the temporal order of these assemblages. We could then resort to the indirect method known as SERIATION.

Seriation is customarily accomplished by a hand-and-eye, trial-and-error technique. Each assemblage is represented by a strip of paper with segments marked off representing the frequencies, in per cent, of various artifact-types or attributes. The strips are then arranged according to the assumption

that a culture changes gradually and constantly, and that particular styles of pottery or clothing or other manifestations of the culture rise to peaks of popularity at certain times and then give way to other styles.

In the case of our pottery samples, we could assume that usually each collection gave what amounted to a momentary glimpse of the popularity of a pottery style at a particular time and place, since most of the village sites probably were occupied only for a relatively short time. If our survey was sufficiently thorough, the individual collections should cover the whole sequence of pottery development in the region. By arranging our set of "still" pictures in the correct order, we could reconstruct a "motion picture" of this sequence. Proceeding on these assumptions, we classified the 346,099 fragments of pottery we had collected, calculated the popularity of each type at each site and made graphs showing the proportions and kinds of styles in each collection. These graphs were arranged and rearranged in relation to each other until the chronological pattern was discovered."

¹¹ James A. Ford, "Mound Builders of the Mississippi", *Scientific American*, Vol. 186, No. 3 (March 1952), pp. 23-24. Copyright © 1952 by Scientific American, Inc. All rights reserved.

Fig. 1 is a reproduction of a seriation graph, illustrating the results of the above process. A more mechanical method of seriation works like this. We compute an index of likeness for each pair of assemblages, as described above in the discussion of classification of assemblages. This time, however, instead of simply sorting the assemblages into cultures, we arrange them in order of similarity, using a rectangular table, or matrix, such as this one taken from Spaulding.¹²

	A	B	C	D
A	200			
B	170	200		
C	124	154	200	
D	90	120	166	200

Here A, B, C, D represent assemblages and the numbers are the values of the likeness-index; 200, of course, represents perfect likeness, the likeness of an assemblage to itself. A, B, C, and D are arranged and rearranged until the indices decrease uniformly as one moves away from the diagonal of self-relationships, so that each assemblage is most like the one just before and just after it and less like those farther away. To quote Spaulding again :

If the spatial factor is inconsequential, such an arrangement would probably be interpreted as chronological, although there is no information as to whether A is the oldest or the youngest assemblage in the group. [This is true of Ford's paper-strip method as well, of course.] ... The underlying concept is the idea that quantitative changes in culture take the form of smooth unimodal curves, and in one form or another it can be found in archaeological research everywhere.^o

While we are discussing seriation an important CAVEAT must be entered: We have mentioned more than once that an assemblage represents the products of a culture at a particular moment in time. When seriation is being attempted, extreme care must be taken to ensure that this temporal limitation is adhered to as closely as possible. Coe, for example, rather than using different sites as sources for assemblages to be seriated, instead analyzed the contents of each of 37 refuse pits from a single site. He reasoned that the length of time between the opening and closing of a refuse pit would be relatively short, whereas a whole site might have been occupied for many years and the actual cultural changes obscured by the mixing up of the debris from all those years. This in

¹² *Op. cit.*, p. 82.

¹³ *Ibid.* See Driver, *op. cit.*, for a discussion of the historical development of seriation.

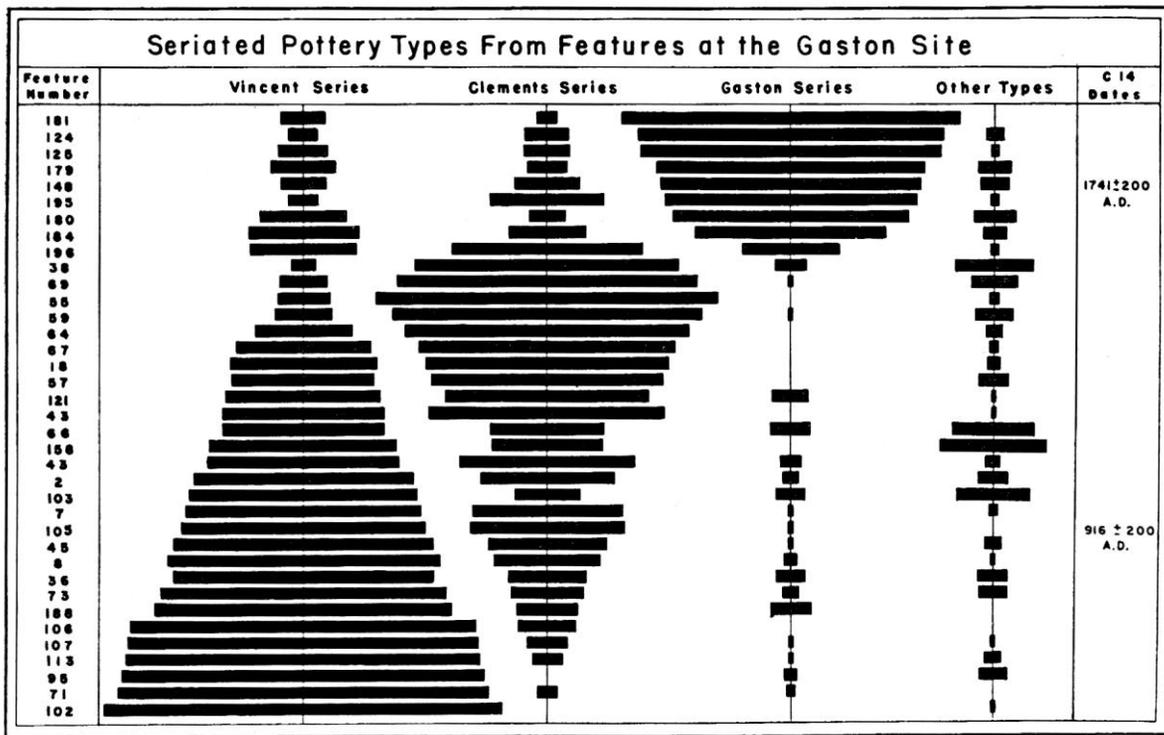


Figure 1. A seriation graph. The features were all refuse pits, and a series is a type or, better, a type of types, of pottery. The lengths of the bars after a particular feature number represent the frequencies (percentages) of each of the types in the assemblage of potsherds from that feature. Using the radiocarbon (C 14) dates to determine which end of the graph is earlier, it can be inferred that: (1) The Vincent Series was extremely popular at the beginning of the period represented, and its popularity diminished steadily to the end. (2) The Clements Series was first made very early in the period, became gradually more and more popular, and then its popularity declined rather abruptly near the end. (3) The Gaston Series came in sometime before 1741, and became very popular quite rapidly, in effect displacing the Clements Series, which had in turn displaced the Vincent Series. The indications of Gaston pottery in early pits are probably due to accidental intrusions of sherds into old pits. (Reproduced from Joffre Lanning Coe, *The Formative Cultures of the Carolina Piedmont* [= *Transactions of the American Philosophical Society* [N.S.], Volume 54, Part 5] [1964], p. 100. Copyright 1964 by The American Philosophical Society. Used by permission.)

turn could have resulted in the total assemblage being misplaced in the temporal sequence, and the history of the culture being misconstrued."

Assuming that the data are all right, then, we can use a computer not only to calculate the values of the index of likeness for the pair assemblages, but also to order the assemblages to produce the closest possible approximation to a matrix of the sort shown above. Indeed, a program has been written by Marcia and Robert Ascher¹⁵ which does both, printing out the resulting table and identifying the assemblages which don't fit as well as they should. Computation of an 8 x 8 matrix took "less than one minute".¹⁶

¹⁴ Joffre Lanning Coe, *The Formative Cultures of the Carolina Piedmont* (= *Transactions of the American Philosophical Society* [N.S.], Volume 54, Part 5) (1964), pp. 100, 107. Fig. 1 is a reproduction of Coe's seriation diagram.

¹⁵ Marcia Ascher and Robert Ascher, "Chronological Ordering by Computer", *American Anthropologist*, Vol. 65 (1963), pp. 1045-1052.

¹⁶ *Ibid.*, p. 1050. Since the above was written, another, more sophisticated seriation program has been developed. This program utilizes iterative procedures to avoid input-order bias, a problem

With seriation, then, we complete the roster of the analytical techniques of scientific archaeology, at least as far as computers have been used, up to the present." We are in a position to make statements such as "these flint-knappers or potters or whatever had such-and-such ideas about flint-knapping or pottery, and they had these ideas because they belonged to period P 10 in the historical development of culture C 1".

The method outlined below can be used as a substitute for seriation, that is to reconstruct an historical sequence of cultural changes. When used for that purpose, it

which the Aschers' method does not solve. The authors of this new program do not give running times in their report, but it can be safely inferred that at least several minutes are involved. Richard S. Kuzara, George R. Mead, and Keith A. Dixon, "Seriation of Anthropological Data: A Computer Program for Matrix-Ordering", *American Anthropologist*, Vol. 68 (1966), pp. 1442-1455.

¹⁷ I have not attempted to discuss what might be called "socio-economic reconstructions", based on archaeological data. An extremely interesting example of the latter is described in J. C. Gardin, "Reconstructing an Economic Network in the Ancient East with the Aid of a Computer", in Hymes, *op. cit.*, pp. 377-391.

differs from seriation in that one only asks whether or not a given type of attribute occurs in an assemblage, rather than asking for the FREQUENCY of the type or attribute."

Besides reconstructing a sequence of changes, this method can also, under certain circumstances which I will discuss later, distinguish developmental, orderly change from random or accidental change.

For the first task, that of simply reconstructing the sequence of events, suppose that we have a large number of assemblages. Suppose, further, that these assemblages are relatively small and yet each contains a wide range of types of artifacts; in many cases there are only one or two of each type in any assemblage. Frequency distributions, then, are meaningless. (One item in fifty is not, properly speaking, 2 % of the assemblage; it is either much more significant than that or else it is utterly insignificant.) Now we want to find the temporal order in which these assemblages were laid down and, through that, describe the historical development of the culture.

Just as we made the assumption for seriation that the popularity of a type will make a unimodal curve, so we make an assumption here: We assume that once a new type or attribute is adopted, it is not lost again during the time covered by our collections; or, once a type present at the beginning of our sequence is lost, it will not reappear. If our sequence is not too long, this assumption will hold for most of the types in question. Such exceptions as there are should not throw our sequence off.

Let us imagine, then, that we have five assemblages, A, B, C, D, E.

A B C D E

A given trait or type or feature is found in A, C, and I and not in B or E:

	A	B	C	D	E
trait i	X	O	X	X	O

Now that cannot be the correct temporal order of the assemblages, according to our assumption, for that order entails that trait i was lost after the period of A, regained after the period of B, and lost again after D. To put it another way, **B** and E must both be earlier than A, C, or D, or they must both be later. Let us temporarily assume the former: **B** and E are earlier than A, C, D, and trait i has been ADOPTED in the course of the sequence:

¹⁸ Dempsey and Baumhoff have reported a method of establishing historical sequences on the basis of presences and absences of artifact types. The method reported here bears a certain similarity to theirs, although it was developed independently. See Paul Dempsey and Martin Baumhoff, "The Statistical Use of Artifact Distributions to Establish Chronological Sequence", *American Antiquity*, Vol. 28, No. 4 (1963), pp. 496-509.

	B	E	A	C	D
trait i	O	O	X	X	X

Now we look at the distribution of another trait, j, according to our present arrangement of assemblages:

	B	E	A	C	D
trait i	O	O	X	X	X
trait j	O	O	X	O	X

We see that the assemblages are not yet in the correct order; we must reverse A and C:

	B	E	C	A	D
trait i	O	O	X	X	X
trait j	O	O	O	X	X

Suppose when we add our next trait, K, it does this:

	B	E	C	A	D
trait i	O	O	O	X	X
trait j	O	O	O	X	X
trait k	O	X	O	O	O

Again, we must reorder the assemblages and we can reorder them, but this time, if i and j were ADOPTED during the sequence, k must be a trait which was LOST during the sequence:

	E	B	C	A	D
trait i	O	O	X	X	X
trait j	O	O	O	X	X
trait k	X	O	O	O	O

Carrying this process out, we find eight patterns of trait distribution that conform to the assemblage sequence E B C A D. These patterns, then, are

	E	B	C	A	D
k	X	O	O	O	O
p	O	X	X	X	X
m	X	X	O	O	O
i	O	O	X	X	X
o	X	X	X	O	O
j	O	O	O	X	X
l	X	X	X	X	O
n	O	O	O	O	X

Supposedly, then, if our assumption of one-way loss or gain holds good for every trait and there have been no intrusions into our assemblages, these are the only patterns that will be found. That is to say, there will be no traits, for example, with these distributions:

q	O	X	O	O	X
r	X	X	O	X	X

Of course our assumption of one-way loss or gain may not hold for every trait, and there almost certainly will be

sources of error such as intrusive artifacts in an assemblage, mistakes of omission in the inventory of assemblages, and the like. Thus we have to order and re-order the assemblages until we get as few errors as possible, until we get that ordering of assemblages which approximates, as closely as possible with the data we have, the ideal ordering.

Then we proceed to make inferences about history, as follows: "At the beginning of this sequence, culture C 1 had traits k, m, o, l (and others with those same or similar distribution-patterns). This beginning period produced, and is represented by, assemblage E. A little later on, k had been lost and p had been adopted, along with other traits similarly distributed, and assemblage B was produced. Then m was lost and i was adopted, producing assemblage C. . ." and so on. We are reconstructing the culture history, and thus explaining why the producers of the artifacts in each assemblage had it in their minds to produce what they produced.

Incidentally, it should be noted that this method, like seriation, does not establish which end of the sequence is the earliest. That must be determined by other means.

When we set up an example with five assemblages and a dozen or so artifact types it is very simple to find the best arrangement. When the numbers of assemblages and traits get much larger, and the frequency of not-perfectly-correct rows goes up accordingly, a computer becomes essential.

The type case for this is the study of a Trinidadian village I mentioned before." The data for this study included not only household artifacts, like tables, dishes, outbuildings, and so on, but also verbatim responses to questions asked of the male and female heads of the household. These responses were coded into types or classes just as were the artifacts, so that each household was inventoried for every response-type and for every artifact-type. We have already adopted the practice of referring to both response-types and artifact-types as TRAITS. The results of the inventory of one household constitute an ASSEMBLAGE, in exactly the same sense as the excavated contents of a refuse pit constitute an assemblage.

With 28 assemblages and 3,666 traits from my Trinidadian study, my programmer" and I developed a computer routine which rearranged the assemblages to produce the best approximation to the ideal. The corn-

puter was fed a deck containing a card for each trait, with the assemblages in alphabetical order; and one additional card giving a suggested sequence of assemblages. The computer then determined which traits would be being adopted during the sequence and which being lost, if that suggested sequence were the right one. Then it determined what sequence of assemblages best fitted the TRAITS, assuming that those adoption and loss designations were correct. Then that new sequence was used as the suggested sequence for another cycle of the same routine, and this was continued until the same sequence of assemblages appeared twice in a row. The computer then calculated the cut-off point for each row (the place where, theoretically, the trait was adopted or lost), and punched a new card for each trait, this time with the assemblages (households) in the correct order, from the most conservative or old-fashioned to the most progressive or modern.

As described up to this point, the method does exactly what seriation does, that is it shows the history of the culture of a people.

Now many of the changes that make up a people's cultural history are accidental, the result of random chance. Other changes are idiosyncratic, the result of the peculiarities of a single mind. Still others are, from the point of view of the people in question, involuntary — various things happen to their culture as a result of natural forces or the actions of alien peoples. For example a pottery-style may be adopted for the first time simply because an alien people who make pottery in that style have just moved into the neighborhood. The loss of a trait is more difficult to explain on extrinsic grounds, except when the source of raw material is cut off, by an enemy or by nature. "Boredom" seems to be the explanation offered for much of cultural loss, but that doesn't explain why a people gets bored with one trait and not with another they've had longer.

I would claim, however, that there are non-random, non-idiosyncratic, voluntary, INTRINSIC causes of cultural change which are fully as important as the involuntary or extrinsic causes, and much more rewarding to study. I say that, given a certain habitat, and given certain alien neighbors with alien ideas available for adoption, a culture will tend to borrow and lose traits voluntarily and systematically, in a certain sequence. That this is so follows from the general principle that a human culture is a functionally integrated whole, as I have shown in detail elsewhere.²¹

Now because chance factors, idiosyncratic behaviors, and extrinsic causes operate to obscure what I call the "NATURAL order of cultural adoption and loss" of a

²¹ Cloak, *loc. cit.*, pp. 7-45.

¹⁹ A more detailed account of the method, along with its theoretical rationale and a description of the field work methods, is contained in F. T. Cloak, Jr., *A Natural Order of Cultural Adoption and Loss in Trinidad* (= *Working Papers in Methodology*, No. 1, Institute for Research in Social Science, University of North Carolina) (1967).

²⁰ Mr. Robert Schacht, University of Wisconsin Computing Center, Madison, Wisconsin.

culture, it is necessary to use probabilistic techniques for establishing the existence of that order.

One essential point should be interpolated here: If we wish to test for the existence of a natural order or sequence, the assemblages must not only be instantaneous, they must be CONTEMPORANEOUS. It may seem paradoxical to use contemporaneous assemblages for the study of historical change, but there is a good reason. If several assemblages were the product of the very same people at several points in time, the trait-differences among them could be due to chance invention or changes in extrinsic causes as well as to the intrinsic natural order. If, however, the assemblages are contemporaneous, we can infer that they were subject to approximately identical extrinsic forces; thus any regular differences in them are due to their being at different stages of the natural order of their common culture. Differences which are due to idiosyncratic, extrinsic, and chance factors will show up as irregular; i.e., will appear as errors, causing "incorrect" distribution-patterns for various traits.

In order to be certain I had contemporary assemblages,

and to get a large amount of qualitative data for each, I used my Trinidadian ethnographic materials. Each "assemblage", then, consisted of responses to questionnaires and observations of artifacts in a single household. The results were processed as described above, producing that order of households which in turn produced the fewest errors in the trait-distributions. Then Monte Carlo methods were invoked to find the probability of coming as close to the perfect natural order pattern with data generated by chance alone as I had come with my field work data. This was done as follows:

The data, when arrayed, form a binary matrix of X's and O's, 28 household-assemblages wide by 3,666 traits long. Forty sample matrices were drawn at random from this empirical matrix, each 12 wide by 366 long. Each of these matrices was processed just as the big matrix had been, and then its DISTANCE from the ideal processed matrix, the one with only perfect rows, was measured. The distance was measured by counting the number of times an X was on the wrong side of an O in each row, and then totalling these counts for all the rows. To

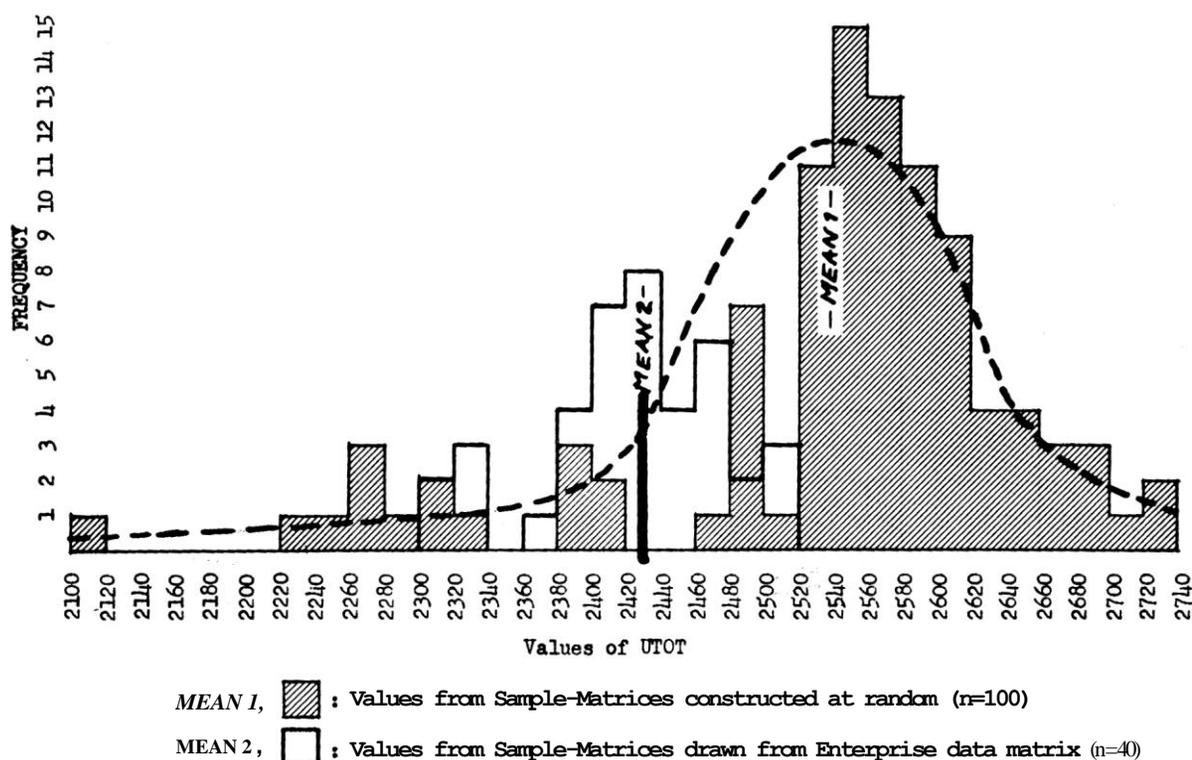


Figure 2. Graph used to estimate the probability that a natural order principle is operating. "UTOT" is the name, in FORTRAN, for the distance from perfection of a matrix (see text for definition). Mean 2 is the average of values from the data gathered in Trinidad ("Enterprise" is the name of the village). It cuts off about 16 % of the roughly normal curve drawn over the shaded histogram, which represents values generated at random. (From F. T. Cloak, Jr., *A Natural Order of Cultural Adoption and Loss in Trinidad* [= *Working Papers in Methodology*, No. 1, Institute for Research in Social Science, University of North Carolina] [1967], p. 141.)

illustrate, suppose we are dealing with a matrix eight columns wide. A correct row for an adopted trait with three X's, then, would look like this :

O O O O O X X X

If, instead, it was incorrect and looked like this

O X O O O X O X

we find that the first X is on the wrong side of four O's and the second X is on the wrong side of one O, the third X being all right. Four for the first X plus one for the second X gives a count of five for that particular row. Adding up the count of each and every row gives the total distance from perfection for the whole matrix. Adding up the distances of the 40 sample matrices and dividing by 40 gives the average distance from perfection.

Then 100 little matrices of the same dimensions as the sample matrices — 12 x 366 — were GENERATED at random by the computer. These generated matrices were processed and THEIR distances from perfection measured. The distribution of the distances of the generated matrices appeared to be roughly normal (Fig. 2). The average distance of the sample matrices cut off about 16 % of that normal curve; this indicates that one would get as close as I did to a perfect natural order pattern about 16 % of the time if one's data were generated at random rather than being controlled by some directional principle.

Given the number of sources of error, both in the extrinsic and chance factors varying the behavior of the people and in the usual mistakes made in data-collection and interpretation, I was willing to accept that 16 % probability level and state that there is indeed a natural order of cultural adoption and loss helping to determine the direction of culture history in this particular case. I would go further and claim generality for that principle, subject to further testing in other parts of the world.

One serious drawback to my study, or any study using ethnographic materials, is its lack of DEPTH. The most conservative household and the most progressive household of my sample of 28 have between them a cultural gap representing probably no more than 65 years of cultural change. What would be ideal, now, would be for me to fly back in time to a point where the most conservative household of today would be the most progressive, in other words about 65 years. Then I could repeat my study, and thus double the depth of the known natural order. I could then jump back another 65 years and do it again, and so on, thus both learning the cultural history of the people and sorting its intrinsic events from its extrinsic and fortuitous events.

Archaeology, in principle, gives us the means of doing just this sort of time-hopping. We could carry out my proposal if, for instance, we could study a city-site that had been inhabited continuously for several hundred years. This site would further have to have clear markers so that the rubbish from each house and from each period in each house could be treated as a separate assemblage. These assemblages would then be processed, as outlined above, in separate contemporaneous batches.

Perhaps, on the other hand, some of the methods we have described — seriation, automatic classification, and the like — can be combined with the method just outlined to get the same sort of results from less ideal archaeological deposits; or perhaps entirely new concepts and methods will be developed.

At any rate, it seems quite likely to me that archaeology, making a full use of its new tool, the computer, will soon move to still another level of humanistic concern and start asking not only "What was in the producer's mind?", and "How did it get there?" but also "By what processes did the culture that put it there come to have it?"