

March 5, 1978

State Univ. of N.Y. at Binghamton
Binghamton, NY 13901

RECEIVED
MAR 13 1978
UNIVERSITY OF CALIFORNIA
LIBRARY

AN ARTIFICIAL INTELLIGENCE DIRECTED ANALYSIS
OF REGIONAL SURVEY DATA IN ARCHAEOLOGY

The objectives of the project is to apply AI strategies and techniques to an existing data set pertaining to early farming villages in Calabria (southern Italy), in order to make inferences about the evolution of human populations and their cultural and biological adaptations. As the analysis has an important spatial dimension, we will integrate methods of spatial pattern recognition with heuristic procedures derived from the experience of some of the leading investigators in the field.

The project is directly related to work that Ammerman and Cavalli-Sforza (1971; 1973) have been doing on the problem of the spread of early farming in Europe over the last six years. The shift from subsistence economies based on hunting and gathering to those based on food production (agriculture) involved substantial demographic changes which have implications for the biological adaptations of human populations and their population genetics. Europe represents one of the best documented areas of the world for this line of investigation.

One aspect of this work was a field project in Calabria concerned with describing patterns of population growth among early farming groups and the density of early farming villages. Five seasons of survey work in Calabria have led to the location of some 450 prehistoric sites: with over 200 sites (dating to the period from 5,000 B.C. to 2,000 B.C.) providing information on the question of settlement patterns among early farming groups. A major season of excavations (supported by a grant from the NSF) was also conducted this past year (1977) at two early farming (Neolithic) sites in Calabria. For each of the sites located during the survey, information was systematically collected on a wide range of variables: site location, size, shape, collection strategy, artifacts recovered, soils, water resources and so forth. To handle the large amounts of data collected, data files and basic programs for listing, sorting and displaying data were developed by Juliana Hwang, a computer programmer in the Department of Genetics at Stanford, working in collaboration with Ammerman. One of the immediate aims was to develop a package of programs so that the information collected during a given field season could be listed and displayed in convenient form so that it could be used in planning the strategies for subsequent field seasons. This turned out to be highly productive and represents one of the first successful attempts at developing a sequential research design in archaeological survey work. There was also the long term aim (when the data sets became more complete) of developing an AI directed analysis of spatial patterns as they relate to the interactions of cultural and ecological variables.

The project is enabled by the possibility of collaboration with Keith Kintigh who has the necessary computer science expertise and a long standing interest in the use of artificial intelligence techniques in dealing with archaeological problems. This academic year he is working at the SUNY-Binghamton Computer Center while on leave from the Ph.D program in anthropology at the University of Michigan. While an undergraduate at Stanford, he became interested in the application of computer science (and especially artificial intelligence) to archaeology and he went on to complete a MS in Computer Science at Stanford before starting his graduate work in archaeology. At Stanford, he worked with Bruce Buchanan toward developing an AI application to archaeology.

In the course of the proposed project, we plan to enlist the expert assistance of several interested archaeologists on the SUNY campus and at the University of Michigan, including Henry Wright, Jeffrey Parsons, Kent Flannery, Robert Whallon, Charles Redman and William Isbell. We anticipate that these contacts will prove invaluable in formulating our strategics and heuristics.

The central problem in the study of prehistoric settlement patterns (and indeed in much of archaeology) is one of pattern recognition. Whereas archaeologists have been able to collect large amounts of good survey data, our methods of interpretation are far from adequate. For example, Jeffrey Parsons' Valley of Mexico Survey (Parsons 1971; 1974) is considered to be one of the best developed projects concerned with the study of settlement patterns. While Parsons has made substantial progress in developing a conceptual framework for the analysis of his data, it is fair to say that his often studied data have yet to be fully or adequately exploited. In addition, substantial progress has recently been made on the general application of statistical methods to the analysis of archaeological data (e.g. Doran and Hodson 1975; Hodder and Orton 1976; Flannery 1976). We thus feel that the problems are well enough defined and basic analytical techniques sufficiently well developed to attempt an AI approach. We are particularly optimistic because many of the strategics archaeologists traditionally employ are mainly limited by the fact that they are so time consuming to carry out manually, yet they are too complex to reduce to standard kinds of statistical treatments. We are faced with a complex problem of recognizing various sorts of spatial patterns and trying to relate them to a number of interrelated cultural and environmental variables. Yet we are saved from the combinatorial morass in so far as we can use our knowledge of human behavior and the interaction of many of the relevant variables.

We see this project as consisting of several tasks: (1) develop an adequately restricted definition of our problem domain and formulate an outline of the strategy that we will apply; (2) elicit strategics and heuristics from other archaeologists concerned with the interpretation of archaeological survey data; (3) integrate the relevant pattern recognition strategies developed in AI; (4) develop heuristic methods by individually testing the methods against a trial data set.

We have rather modest goals: we are proposing a pilot project to examine the feasibility of our goals and to develop some heuristic strategies. Already this year, we have assembled many of the elements of a trial strategy. The data base for the required experimentation already exists in a convenient representation on the SUMEX system. As our computer effort will be primarily devoted to program debugging and small scale tests, mainly at off-peak hours, we foresee no large scale demand for machine resources. The success of this pilot project would lead us to a longer term, more intensive project probably including Keith Kintigh's doctoral thesis.

Bibliography

- Ammerman, A.J. and L.L. Cavalli-Sforza
1971 Measuring the Rate of Spread of Early Farming in Europe.
Man 6: 674-688.
- Ammerman, A.J. and L.L. Cavalli-Sforza
1973 A Population Model for the Diffusion of Early Farming in
Europe. In The Explanation of Culture Change. Ed. C. Renfrew.
London: Duckworth.
- Doran, J.E. and F.R. Hodson
1975 Mathematics and Computers in Archaeology. Edinburgh:
Edinburgh University Press.
- Flannery, K.V. (ed.)
1976 The Early Mesoamerican Village. New York: Academic Press.
- Hodder, I. and C. Orton
1976 Spatial Analysis in Archaeology. Cambridge: Cambridge
University Press.
- Parsons, J.R.
1971 Prehistoric Settlement Patterns in the Texcoco Region,
Mexico. Memoirs no. 3. Museum of Anthropology,
University of Michigan.
- Parsons, J.R.
1974 The Development of a Prehistoric Complex Society: a Regional
Perspective from the Valley of Mexico. Journal of Field
Archaeology 1: 81-103.

Ammerman and Kintigh

An Artificial Intelligence Directed Analysis of Regional Survey
Data in Archaeology

The entries below follow the sequence of the SUMEX-AIM resource Questionnaire. A general description of the proposed research and how it fits in with our previous experience is provided in a separate two page proposal.

- A1 See proposal.
- A2 The research does not require support outside of communication costs which we should be able to raise support for. Both Ammerman and Kintigh have positions and other personnel is not required. Kintigh has his own computer terminal at home. We have recently asked for a small amount of support towards communication costs in a grant proposal to the NSF (see attached budget). The awarding of this grant is not contingent on SUMEX-AIM availability.
- A3 The research is closely related to the AI approach of SUMEX-AIM; see proposal.
- B1 Yes, they would have wide application. There are now many regional survey projects in archaeology and the full exploitation of these data sets in terms of analysis would be put on a new footing by the programs we would like to develop.
- B2 Various programs already exist on SUMEX-AIM (e.g. Calabria data files, programs for permutation sorts and graphic display using these files, basic statistics programs for archaeological data). Kintigh has acquired many of the programs for data analysis in archaeology and is familiar with the programs in use at the University of Michigan.
- B3 We anticipate no difficulties in this area and plan to make our programs available to other workers in the field.
- B4 Yes.
- B5 Yes. We appear to be the first investigators starting to explore AI directed analysis in our field.
- C1 Until January 1978, Ammerman was using the SUMEX facility when he was at Stanford. We have access to the SUNY Binghamton computer facility -- Kintigh is, in fact, on the research staff of the computer center -- but the system is not suitable for the research we have in mind. The system is not interactive for research purposes.

- C2 Lisp or Lisp related languages. The research should not require the addition of other system programs or languages.
- C3 Our CPU requirements should be fairly modest, since our main work will be devoted to program development. It is hard to be specific here without knowing more about how the SUMEX system works (this could be made more concrete by means of a discussion between Kintigh and the people at Stanford). Our disk space needs would be about 150 pages. We would plan on using the system during off-peak periods (mainly 16:00 - 21:00 PST which is late evening on the east coast when telephone costs would be cheaper for us). Kintigh has his own terminal. Our expected connection time would probably average about 10 hours per week. We would be seasonal users in the sense that there would be active periods of use and also inactive periods (when we are doing field work).
- C4 TYMNET: Syracuse, New York. Or Corning, if available. Kintigh has his own terminal. We have started looking into support for communication costs (telephone connection to TYMNET node) and should be able to raise the amounts needed through a combination of local sources.
- C5 See proposal.

BUDGET DETAIL
September 1, 1978 - August 31, 1979

E. MATERIALS AND SUPPLIES

Small excavation equipment	350
Plastic sheeting	200
Bags & boxes for storage	350
Photo supplies	200
Drawing supplies	150
Sample containers	<u>100</u>

TOTAL MATERIALS AND SUPPLIES 1,350

L. OTHER DIRECT COSTS

1. Maintenance allowances for specialist (\$500/mo)

<u>Specialist</u>	<u>Position</u>	<u>Months</u>	<u>Total Allowance</u>
Aldridge	Assistant Director	4	2,000
Vacant	Site Supervisor	3	1,500
Carrara	Processing	3	1,500
Bartlett	Magnetometer Survey	1	500
Rommelzwoal	Soils Specialist	2	1,000
Wyonastra	Pollen Specialist	1	750*
<u>Diamond</u>	<u>Obsidian Specialist</u>	<u>2</u>	<u>1,000</u>
(*includes transport of boring equipment)			8,250

2. Maintenance allowances for students (\$250/mo)

7 students for 3 months 5,250

3. Workmen (\$20/day for 100 days) 2,000

4. Land rent 450

Magnetometer equipment and insurance 250

Shipment of samples 200

Processing of samples (C14, NAA Obsidian, Obsidian hydration) 400

5. Long distance connection for computer use 300

8,850

TOTAL OTHER DIRECT COSTS \$17,100