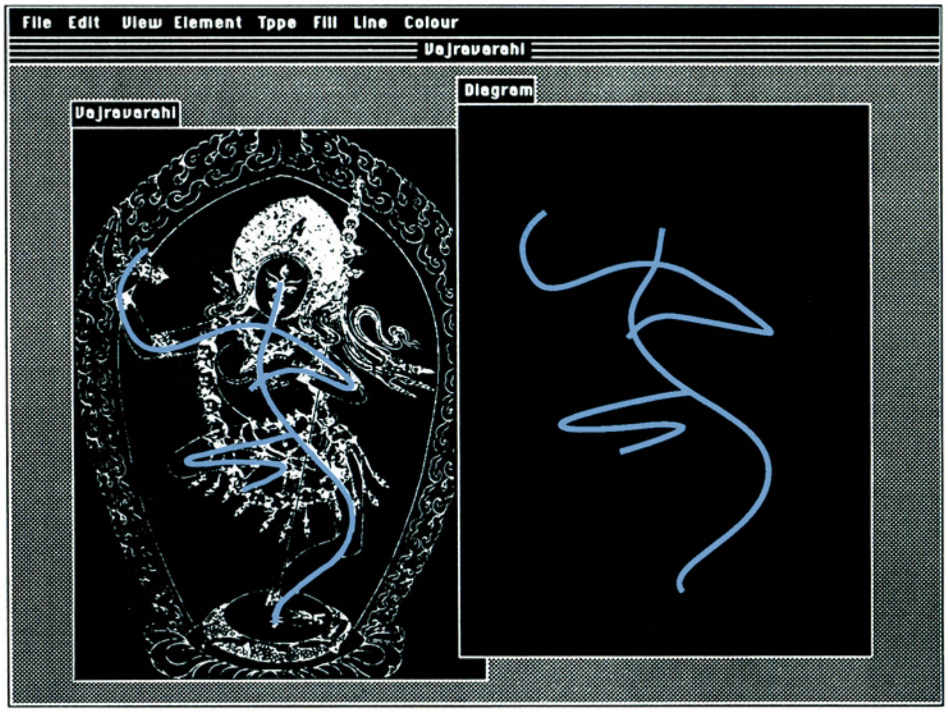
Archaeology and the Information Age

A global perspective

Edited by Paul Reilly and Sebastian Rahtz



ROUTLEDGE

ONE WORLD ARCHAEOLOGY

21

ARCHAEOLOGY AND THE INFORMATION AGE

ONE WORLD ARCHAEOLOGY

Animals into art 1H.Morphy (ed.)

Archaeological approaches to cultural identity
S.J.Shennan (ed.)

Archaeological heritage management in the modern world
H.F.Cleere (ed.)

Centre and periphery: comparative studies in archaeology
T.C.Champion (ed.)

Conflict in the archaeology of living traditions
R.Layton (ed.)

Domination and resistance
D.Miller *et al*. (eds)

The excluded past: archaeology in education
P.Stone & R.Mackenzie (eds)

Food, metal and towns in the archaeology of Africa
B.Andah *et al.* (eds)

Foraging and farming: the evolution of plant exploitation
D.R.Harris & G.C.Hillman

From the Baltic to the Black Sea: studies in medieval archaeology
D.Austin & L.Alcock (eds)

Hunters of the recent past
L.B.Davis & B.O.K.Reeves (eds)

The meanings of things: material culture and symbolic expression
I.Hodder (ed.)

The origins of human behaviour
R.A.Foley (ed.)

The politics of the past
P.Gathercole & D.Lowenthal (eds)

Signifying animals: human meaning in the natural world
R.G.Willis (ed.)

State and society: the emergence and development of social hierarchy and political centralization
J.Gledhill *et al.* (eds)

The walking larder: patterns of domestication, pastoralism, and predation
J.Clutton-Brock (ed.)

What is an animal?
T.Ingold (ed.)

What's new? A closer look at the process of innovation
S.E.Van der Leeuw & R.Torrence (eds)

Who needs the past? Indigenous values and archaeology
R.Layton (ed.)

ARCHAEOLOGY AND THE INFORMATION AGE

A global perspective

Edited by

Paul Reilly
IBM UK Scientific Centre, Winchester

Sebastian Rahtz
Department of Archaeology, University of Southampton

London and New York

First published 1992
by Routledge
11 New Fetter Lane, London EC4P 4EE

This edition published in the Taylor & Francis e-Library, 2005.

Simultaneously published in the USA and Canada
by Routledge
a division of Routledge, Chapman and Hall, Inc.
29 West 35th Street, New York, NY 10001

British Library Cataloguing in Publication Data
Archaeology and the Information Age: a global perspective. —(One world
archaeology) I. Reilly, Paul II. Rahtz, Sebastian 930.10285

Library of Congress Cataloguing in Publication Data

Also available

ISBN 0-203-16834-8 Master e-book ISBN

ISBN 0-203-26356-1 (Adobe eReader Format)
ISBN 0-415-07858-X (Print Edition)

List of contributors

Tim Allen, Oxford Archaeological Unit, Oxford, UK.
Martin Biskowski, Department of Anthropology, University of California Los Angeles, USA.
Gill Chapman, Design Department, Sheffield City Polytechnic, UK.
Costis Dallas, Benaki Museum, Athens, Greece.
Ken Delooze, Computer-Aided Engineering Centre, North Cheshire College, Warrington, UK.
Pavel Dolukhanov, Department of Archaeology, University of Newcastle-upon-Tyne, UK.
Mike Fletcher, Mathematics Department, Staffordshire Polytechnic, Stafford, UK.
Wendy Hall, Department of Electronics and Computer Science, University of Southampton, UK.
Trevor Harris, Department of Geography, University of West Virginia, USA.
Karega-Munene, Department of History, University of Nairobi, Nairobi, Kenya.
Todd Koetje, Department of Anthropology, State University of New York at Binghamton, USA.
Martin Kokonya, National Museums of Kenya, Nairobi, Kenya.
Marie-Salomé Lagrange, Centre National de la Recherche Scientifique, Paris, France.
Gary Lock, Institute of Archaeology, University of Oxford, UK.
Ranjit Makkuni, System Sciences Laboratory, Xerox Palo Alto Research Center, California, USA.
Arkadiusz Marciniak, Institute of Prehistory, Adam Mickiewicz University, Poznań, Poland.
Roger Martlew, Department of Extra-mural Studies, University of Leeds, UK.
Maria Meneses, Department of Archaeology, Eduardo Mondlane Universty, Maputo, Mozambique.
Brian Molyneaux, Department of Archaeology, University of Southampton, UK.
Akifumi Oikawa, National Institute for Education, Tokyo, Japan.
Włodzimierz Rączkowski, Institute of Prehistory, Adam Mickiewicz University, Poznań, Poland.
Sebastian Rahtz, Department of Archaeology, University of Southampton, UK.
Jean-Aimé Rakatoarisoa, Musée d’Art et d’Archéologie, Université d’Antananarivo, Madagascar.
Paul Reilly, IBM UK Scientific Centre, Winchester, UK.
Stephen Shennan, Department of Archaeology, University of Southampton, UK.
Paul Sinclair, Department of Archaeology, University of Uppsala, Sweden.
Dick Spicer, School of Computing, Staffordshire Polytechnic, Stafford, UK.
Arthur Stutt, Department of Archaeology, University of Southampton, UK.
Attila Suhajda, Independent Researcher, Szekesfehervar, Hungary.
Victor Trifonov, Institute of Archaeology, Academy of Sciences, Leningrad, USSR.
Lana Troy, Department of Egyptology, University of Uppsala, Sweden.
Michael Trueman, Lancaster University Archaeological Unit, UK.
Vanda Vitali, Department of Metallurgy and Materials Science, University of Toronto, Canada.
John Wilcock, School of Computing, Staffordshire Polytechnic, Stafford, UK.
Jason Wood, Lancaster University Archaeological Unit, UK.

Foreword

This book is the first in the *One World Archaeology* series to derive from the Second World Archaeological Congress (WAC 2), held in Barquisimeto, Venezuela, in September 1990. Despite many organizational problems (Fforde 1991, p. 6), over 600 people attended the Inaugural Session of WAC 2, with more than 450 participants from 35 countries taking part in academic sessions, and additional contributions being read on behalf of many others who were unable to attend in person.

True to the aims and spirit of WAC 1 over three quarters of the participants came from the so-called Third and Fourth Worlds (see Fforde 1991, p. 7 for details) and the academics came not only from archaeology and anthropology but from a host of related disciplines.

WAC 2 continued the tradition of effectively addressing world archaeology in its widest sense. Central to a world archaeological approach is the investigation not only of how people lived in the past but also of how and why those changes took place which resulted in the forms of society and culture which exist today. Contrary to popular belief, and the archaeology of some 25 years ago, world archaeology is much more than the mere recording of specific historical events, embracing as it does the study of social and cultural change in its entirety.

Like its predecessor, this Congress was organized around major themes. Several of these themes were based on the discussion of full-length papers which had been circulated previously—or were available to be read at the Congress itself—to all those who had indicated a special interest in them. Other sessions, including one dealing with an area of specialization defined by period and geography and another focusing on a particular form of agricultural practice, were based on oral addresses, or a combination of precirculated papers and lectures.

Archaeology and the Information Age: a global perspective results from discussions over five mornings, organized by Paul Reilly and Sebastian Rahtz, of the three volumes of precirculated papers which they had published prior to September 1990. Following the Congress—held only one year ago—the two organizers selected certain of the papers for inclusion as chapters in this volume, and authors were given time for revision and updating of their contributions in the light of discussions at WAC 2. The speed of publication of *Archaeology and the Information Age* represents considerable effort and careful planning since 1988 when this WAC theme was first envisaged.

I should admit at this point that, as an academic educated in the middle and late 1950s within anthropology and archaeology in the UK, I was—until the above decision in 1988—almost entirely computer-illiterate. To my teenager household I was an object of despair as I struggled to get on terms with word processing; to my Departmental colleagues I was clearly of a different generation. Yet, my first years of research were coterminous with the 'new archaeology' and its emphasis on quantitative analysis and model building; indeed, several of my earliest published assertions were supported by chi-squared tests and similar statistics. Subsequently— through my archaeological professional career—it has seemed obvious to me that everyone should be as precise as possible when using terms such as 'trend', 'association', 'expectation'. I have also been intellectually curious and intrigued by some of the research projects being undertaken by my colleagues but have often failed to find the time to seek clarification of many of the terms casually used by them in discussion, and have thus felt, in this context, outmoded, ignorant and dated.

In what follows, therefore, my aim is not to review the main themes of *Archaeology and the Information Age*—which have been detailed in the editorial Introduction— but to examine a few of the points which have struck me personally as being of particular note or fascination.

Earlier this year, in Zimbabwe, a postgraduate research student—employed in the museum at Harare—asked my advice regarding his study of animal and human figurines. I found him seated at his (Swedish-derived) personal computer, manipulating his data in a quick and reliable way that would have been quite beyond me when I was facing comparable problems regarding the analysis of figurines in the early 1960s. It is not only that the technology is available but that the user has no fears.

Archaeology and the Information Age is just the book for people such as me. It does not set out to teach how to use particular forms of soft or hardware; it does not even particularly seek to explain terminology (and acronyms). Instead, it launches the reader into the principles involved in, and the results deriving from, applications of quantitative methods and computers to archaeological questions and archaeological situations.

In so doing, the book does in fact clarify many of the terms and concepts which had remained—at least for me—only vaguely intelligible. I am no longer, for example, still wondering whether I had perhaps misheard the words used by my colleagues in their discussion of something apparently referred to as a ‘virtual reality’, nor do I still imagine little people mysteriously inhabiting screens when I hear about pixels, and I no longer shudder at the mention of radiosity techniques. For someone not interested in actually understanding how such techniques work, this book affords a marvellous way to learn—through the context of applications of such techniques.

However, for the archaeologist, *Archaeology and the Information Age* should do much more than clarify. It should shock. This book demonstrates two fundamental points: first, that the spread of technologies in the context of their archaeological applications is likely to be accompanied by major social upheaval—upheaval which may still develop either for the better or worse; and, second, that work and research within the sphere of applications of computer and quantitative methods to archaeological endeavour are already engaged in a fundamental questioning of the principles and concepts which have long underpinned traditional archaeological enquiry and interpretation.

What has taken me by surprise in this book is not so much the realization that some of the computer applications described within its pages may or may not work (now or in the future) and thereby produce new pictures of the past, but that—whether or not these eventuate—I need to reconsider several of my own preconceptions about method and about the nature of the archaeological record. Many of the questions raised in several of the chapters in this book are fundamental ones.

The combination of the social and conceptual impact of the Information Age on the practice of my chosen discipline makes this book essential reading. This same combination also explains why this book sits so happily within the aims of the *One World Archaeology* series: as described within its pages, the new technologies have the potential to make knowledge about the past more readily accessible to all, but there is also the opposite possibility, that archaeological practice may become even further controlled by the few. It is therefore vital that archaeologists should confront what amounts to an ethical question, and do all in their power to ensure the success of the former development, and prevent the latter.

There are also other problems that have to be faced. Even the obviously positive development of the lowering of book prices, through the use of computer typesetting and so on, which would make books such as these available to Third World colleagues, could have some possibly divisive and complex consequences. Uncontrolled production and dissemination of archaeological data and interpretation has at least the potential to threaten standards of content and scholarship. Indeed, the actual difficulties—and time-consuming nature—of electronic type-setting in desktop publishing is usually grossly underestimated and the interests of speed of production too often conflict with such ‘mundane’ editorial aims as correct referencing, deletion of repetition, standardization of spelling and abbreviations, accurate indexing and so on.

However, it is the whole concept and nature of a ‘book’, as we understand such an entity at the present time, which this volume suggests may be outmoded (and see *The meanings of things,* edited by I.Hodder). This is not a matter of end-pockets of books containing fiches, or other such relatively simple ways of disseminating large quantities of data to a relatively small number of readers, but, rather, a new relationship between the written and spoken word, the pictorial image and the reader—the latter being offered alternatives to sequential following of a given text as well as the freedom of selecting levels of information suitable to his or her particular interests and/or level of education.

It is clear from *Archaeology and the Information Age* that the paucity of good and reliable in-depth studies of the reasons for, and the results of, visits to archaeological museums and archaeological sites (and see *The excluded past,* edited by P.Stone & R. MacKenzie) is likely to be even more calamitous in the Information Age than it now is. Current debates about the nature of ‘heritage’—entertainment versus learning, or a combination of the two—have highlighted the deficiencies in our understanding of what constitutes effective site and museum presentation. These debates have also stressed the currently inadequate nature of collaboration between archaeologists, museum curators and display experts. This book, however, makes it clear that the future of public interaction with the past is likely to be qualitatively different from anything that we can currently envisage. Effective computer interactions of the kind described as already in existence in some exceptional museums will undoubtedly force changes on the heritage industry which are really difficult to conceptualize; but they will be nothing to those which archaeologists will have to accept in order to provide the detail and quality of information required for effective interaction with the public, and manipulation of that data by the public itself.

The publication of *Archaeology and the Information Age* coincides with new *Advanced* and *Advanced Supplementary Level* archaeology syllabuses becoming available to schools in the UK, syllabuses which some consider will force change on most first year undergraduate university courses, both in terms of content and teaching methods. This book demonstrates the direction that some of these changes will surely go. Now that excavations may be simulated and different manipulations of the data attempted—without any great prior grasp of computer techniques—so the student (at school or in the university) may learn for him or herself the profound results of an error of judgement in how to excavate—ie. through graphically exemplifying and representing the destructive nature of archaeological excavation—as well as the nature of archaeological decision-making in the context of practicalities such as limited budgets, the nature of available expertise, specialist equipment needs, and so on.

It is indicative of the nature of this book that a reader with no particular interest in, for example, classical Attic stelai, will be gripped by the detailed chapter of nearly fifty pages devoted to them. Of particular interest to me was the clear revelation that computer applications to this archaeological material raised doubts —or at least serious questions—in my own area of research, concerning all existing attempts to recognize compositions within palaeolithic cave art! This was not so much through some wonderful new technique, but because the research presented in this book—through its aim to investigate the specificity and nature of possible mechanisms of communication—made explicit a number of possible ways of recognizing composition which should, also, be testable in the context of the visual arts of other cultures. Although unable to assess the details of the quantitative methods being discussed, I am convinced that the questions of method and principle raised by such quantitative applications offer a new approach—even to the non-numerate— to the analysis of the visual arts, and offer new insights into the way that such arts may act as forms of communication (and see *Animals into art,* edited by H.Morphy).

The chapter on Tibetan *thangka* paintings is equally illuminating. From now on, any of my students who are interested in art and society, or the preservation and presentation of the heritage, will be encouraged to become familiar with at least this chapter in *Archaeology and the Information Age*. As a case study demonstrating the inextricably intertwined nature of archaeological and anthropological research with the politics of cultural identity and cultural cohesion (and see *Archaeological approaches to cultural identity,* edited by S.J.Shennan, and *The politics of the past,* edited by P. Gathercole & D.Lowenthal)—carried out in this case in the context of education and heritage preservation—it is invaluable. As an illuminating example of self-conscious applied action which will, almost inevitably, change a Tibetan tradition into something ‘new’ (and see *Whats new?,* edited by S.E. van der Leeuw & R.Torrence)—but still unmistakably Tibetan—it raises profound questions of principle concerning the role of those involved in heritage preservation.

Questions regarding power and control—through access to the computer hard and softwares needed to carry out the kinds of research described in this book—are central issues in many chapters of *Archaeology and the Information Age*. Indeed, the book’s subtitle—‘a global perspective’—accurately reflects another of its striking features, the first part of the book serving to demonstrate the vastly unequal availability— and consequent differing uses made—of computer technology within archaeological enquiry in different countries of the world (particularly in eastern Europe and Africa).

Such questions of power and control are often also more subtle ones involving control of the human imagination and human critical faculties by the very machines and programs which they have devised. Used without careful consideration, information technology may stifle thought and deaden vitality and heterogeneity. Used with discretion, information technology can be an exciting, and even revolutionary, force which should lead archaeologists to re-examine many accepted preconceptions and unsubstantiated assumptions—the technology remaining at the service of the human masters.

Perhaps one of the best hopes for the positive future development of computer and quantitative applications within archaeology lies with them becoming genuinely available within the so-called Third World. *Archaeology and the Information Age* suggests, but does not consider in detail, that many of the already available computer programs may be applied, and results developed, without any prior knowledge or understanding of the mathematics and engineering involved. If the potential rigidification of thought, through the use of predetermined parameters embedded in programs, is actively counteracted by the different conceptual assumptions and preconceptions of archaeologists deriving from non-western traditions, then the future does indeed appear exciting and challenging.

P.J.Ucko
Southampton

Reference

Fforde, C. 1991. The Second World Archaeological Congress, *World Archaeological Bulletin* 5, 6–10.

Contents

VISUALIZATION

ANALYSIS

COMMUNICATION

List of colour plates
(between pp. 104 and 105)

Preface

Readers should not regard this book as just another catalogue of the wonders of computer technology applied to archaeology, although we hope it will be as stimulating for those already heavily committed to information technology (IT), as for those who are concerned with its theoretical impact in the future, or who are unconvinced of its benefits. This volume records what we regard as an important time in the history of archaeological communication around the world. The technology has reached a stage where it is potentially of use to archaeologists of all kinds—with diverse interests—rather than to just those interested in quantitative analysis, the topic which many archaeologists still associate exclusively with computers. Yet, while it is cheap enough for all countries to be able to afford to use it, the rate of progress each year means that it is not easy to keep up with current work.

A glance at the available literature is sufficient to realize the pace of development within studies of computer applications to archaeology. There are many specialist journal publications. America has produced several newsletters edited by Sylvia Gaines: *Newsletter for Computer Archaeology* and *Advances in Archaeological Computing;* sadly both are now defunct; also from the USA is a newsletter on computing in anthropology *(Computer Applications in Anthropology Newsletter)*. In Europe, *Archeologia e Calcolatori* started publication from Rome in 1990, while specialist journals include *Science and Archaeology*. The University of Århus produces the KARK newsletter; in the UK there is a specialist newsletter, *Archaeological Computing Newsletter*, which has increased in sophistication since its launch in 1985, and the yearly *Computer Applications and Quantitative Methods in Archaeology (CAA)* conferences have produced an important volume of proceedings every year since 1973 (since 1988 they have been published by British Archaeological Reports, Oxford). The journal proceedings of the international *Archaeometry* conferences include papers on computing, as do journals such as *Science and Archaeology* and *Computers and the Humanities*.

Since the excellent survey of Richards & Ryan (1985), there has been no general book in English on computers and archaeology. A bibliography of archaeological computing mainly in the UK is given in Ryan 1988. A bibliography based on Ryan, and updating him, appears as ‘Une bibliographie sur l’application de l’informatique en archéologie’ by D.Arroyo-Bishop & M.T.Lantada Zarzosa, published by the Centre National de la Recherche Scientifique in 1990. Djindjian (1990) has published an interesting bibliography of quantitative applications to archaeology as practised in France. A general volume, *The Humanities Computing Yearbook*, is published periodically by Oxford University Press (Lancashire & McCarty 1988; Lancashire 1991) and includes a section on archaeology.

The present book is not intended to rival, nor does it attempt to synthesize, these works. Rather, it is intended to place earlier work into a wider context, to show what global effect IT is having on archaeology.

This book has its origins in a set of precirculated papers (Reilly & Rahtz 1990a; Reilly & Rahtz 1990b; Reilly & Rahtz 1990c) discussed at the Second World Archaeological Congress (WAC 2) sessions, held in Barquisimeto, Venezuela, in September 1990. When we started to organize these conference sessions in the summer of 1988, we aimed to generate a view on the *impact* of IT on archaeology. We requested papers which showed what benefits IT had contributed to archaeology around the world, and what we could expect in the future. Contributions were sought relating to three major areas, namely: data visualization, formalizing and quantifying archaeological arguments, and information dissemination. We had a wide range of responses, with many radical and thought-provoking contributions, and the sessions in Venezuela were fundamentally stimulating, with fifty communications contributed from seventy participants from many countries (China, Colombia, France, Germany, Greece, Hungary, Italy, Japan, Kenya, Madagascar, Mozambique, Poland, Portugal, Spain, the UK, the USA and the USSR). Papers were presented on *Communication, Data Visualization, Artificial Intelligence and Expert Systems*, and *National Strategies*. Topics ranged from a picture of how archaeological computing, is funded in the UK, a discussion of the usefulness of solid models for visualizing archaeological data, and simulation of archaeological processes for research and teaching, to the computerization of work on the urban origins of east Africa, and the teaching of Tibetan art in a computerized museum.

The history of the World Archaeological Congress (WAC) is well-documented (Ucko 1987) and it would be otiose to expound on its general political background and questions of archaeological theory. What was interesting to us in the creation of this new body in 1986 was the specific aim of *integrating* archaeologists from all countries and races. It is important to stress that the theme of archaeology and *communication*, which underlies this book, was a direct result of the political stance of WAC; rather than hold a forum at which technologists shared tips on hardware and software, we wanted to develop explicit themes to do with how the technology could be used to help bring archaeologists together.

Communication in Archaeology at WAC 2 was an undoubted success. All the participants were unanimous in their vision of a true global discussion of problems which should start straight away. To some extent, of course, there was a danger of simply perpetuating the technology clique, simply expanding it globally, so that archaeologists in even more countries could feel alienated from those few who 'understand computers'. We wanted to stress the use of computers to facilitate the exchange of *ideas* about archaeological concerns. International cooperation depends on international communication; could we hope that a removal of isolation would really promote a democratization of knowledge?

This book offers a further exploration of these issues. The introduction provides general information on the use of IT in archaeology, as a prelude to the detailed studies which follow.

At the Barquisimeto meeting of WAC 2 an important achievement was the agreement by the participants in the *Communication in Archaeology* symposia to propose the establishment of a Special Interest Group (SIG) under the auspices of WAC, to promote research, collaboration and communication between archaeologists with special problems involving IT, at an international level. This proposal was ratified by the WAC Council, and IT was incorporated into WAC's approved areas of study. This book is one outcome of the SIG, and it is to be hoped that by the time of WAC 3, in India in 1994, some of the problems highlighted by the contributions to this book will have been solved.

We are very grateful to all the many individuals and institutions who contributed to the success of the IT sessions at the second WAC. We were assisted in the solicitation and refereeing of papers by Stephen Shennan (UK), Akifumi Oikawa (Japan), Martin Kokonya (Kenya), Paul Sinclair (Sweden) and Fred Plog (USA); the IBM UK Scientific Centre, and the University of Southampton, Department of Electronics and Computer Science, supported the organisation of the sessions, and published the precirculated papers. The organisation of the Congress itself by Milagro Gomez de Blavia was little short of a miracle.

Cressida Fforde's unflagging enthusiasm and hard work smoothed our path in many ways; Leonor Barroca took a central role in solving our translation problems, working on the original *Call for Papers* and several of the final chapters. Gabriela Makkuni, Paula Meneses and Leonor Barroca provided *ad hoc* translation during the Barquisimeto meetings, and the hospitality of the people of Venezuela was magnificent.

The publication of colour figures in this book was only made possible by a generous grant from British Nuclear Fuels plc, to whom we are extremely grateful.

We take this occasion to acknowledge that this publication would never have emerged without the superhuman efforts of Les Carr at the Department of Electronics and Computer Science, University of Southampton.

Our chief debt, however, is owed to Peter Ucko for caring so much about world archaeology and communication, and thus providing the stimulus to make sure that this book reached its readers, and for working so extensively with us on the editing and typesetting.

P.Reilly
(Winchester)
S.Rahtz
(Southampton)
1 September 1991

References

Arroyo-Bishop, D. & M.T.Lantada Zarzosa 1990. *Une bibliographie sur l'application de l'informatique en archéologie.* Paris: Centre National de la Recherche Scientifique.

Djindjian, F. 1990. A select bibliography of French data analysis applications in archaeology. *Science and Archaeology* **32**, 63–68.

Lancashire, I. (ed.) 1991. *Humanities Computing Yearbook.* Oxford: Oxford University Press.

Lancashire, I. & W.McCarty (eds) 1988. *Humanities Computing Yearbook 1988.* Oxford: Clarendon Press.

Reilly, P. & S.Rahtz (eds) 1990a. *Communication in Archaeology: a global view of the impact of information technology. Volume one: data visualization.* Winchester: IBM UK Scientific Centre.

Reilly, P. & S.Rahtz (eds) 1990b. *Communication in Archaeology: a global view of the impact of information technology. Volume two: national strategies, artificial intelligence, communication.* Southampton: University of Southampton.

Reilly, P. & S.Rahtz (eds) 1990c. *Communication in Archaeology: a global view of the impact of information technology. Volume three: late papers.* Winchester: IBM UK Scientific Centre.

Richards, J.D. & N.S.Ryan 1985. *Data Processing in Archaeology.* Cambridge: Cambridge University Press.

Ryan, N.S. 1988. Bibliography of computer applications and quantitative methods. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 1–30. Oxford: British Archaeological Reports.

Ucko, P. 1987. *Academic Freedom and Apartheid: the story of the World Archaeological Congress.* London: Duckworth.

1

Introduction: archaeology and the information age

PAUL REILLY & SEBASTIAN RAHTZ

Computers are a crucial element of modern information technology (IT). They have been with us for most of the latter half of the twentieth century and are proving as influential to modern archaeology as the printing press was to scholarship and science in the fifteenth century. Archaeologists have been lured to them from their inception, mesmerized by their seemingly unlimited potential.

Historically, archaeological computing evolved, mainly in western Europe and north America, from statistical approaches in the late 1950s and early 1960s, and this form of enquiry still commands much attention (eg. Koetje Ch. 13, this volume; Sinclair & Troy Ch. 14, this volume). Perhaps indicative of the established worth and utility of quantitative methods are the number of texts detailing both the basic procedures and the subtle nuances of the quantitative tests and techniques engendered by computers (eg. Doran & Hodson 1975; Orton 1980; Aldenderfer 1987; Shennan 1988; Voorrips & Ottaway 1990).

Unlike many other facets of archaeology which involve specialist procedures, such as those relating to pathology and archaeometry for instance, quantitative methods are rarely consigned to internal exile in the appendices of a report. Nowadays, quantitative methods are so much a part of the archaeological method that they no longer appear in the literature as distinct research papers (eg. Djindjian 1990, p. 63), simply because more material of this kind is being published as integral elements of general archaeological studies.

But where does IT fit into archaeology? Does its importance lie in the fact that it can make it easier and cheaper to promulgate traditional archaeological information, or is the computerization of archaeological knowledge a discipline or sub-discipline in its own right? Does IT in archaeology have its own recognizable goals and achievements, or is it only an ancillary tool which is equally relevant to all branches of archaeology?

In fact, all these positions are to a greater or lesser extent true, depending on what archaeological tradition an individual archaeologist derives from. Examination of computer applications in archaeology in the UK, for example, reveals evidence of support for each of these positions.

Specialists and enthusiasts in computerized archaeological methods began holding regular meetings in Britain in the early 1970s. The annual *Computer Applications in Archaeology* conference—now renamed *Computer Applications and Quantitative Methods in Archaeology,* but still known, anachronistically, as ‘CAA’—has grown from a small club of British scholars and a handful of colleagues from Europe to a unique major international meeting attracting archaeologists from around the world. How deeply the discussions raised at these meetings have penetrated the collective consciousness of archaeologists in the UK, and further afield, however, is debatable. Certainly, several UK field units regularly send representatives to the meetings of the above conference, but in other cases participants are made to take leave to attend—perhaps indicative of the attitude of some individuals in charge of field units and other archaeological organizations. The UK *status quo* is perhaps reflected by its Institute of Field Archaeologists (IFA) which:

> considers it to be important that archaeologists are properly trained in the use of computers and is encouraging the provision of appropriate courses, and the development of computing resources. (Cooper & Richards 1985, p. 10)

Expertise in ‘archaeological’ applications of computers is now recognized as accreditable experience towards fulfilling the IFA’s registered areas of competence requirements for membership of this group. In fact the IFA, working with the Royal Commission on Historic Monuments of England (RCHME), set up a working party which has carried out two surveys looking at what, how and where computerized systems are used in Britain (Richards 1986; Booth, Grant & Richards 1989). The surveys were thought to be necessary in a climate of growing concern about the uncoordinated proliferation of computer-based systems across archaeology in the UK. The working party made recommendations on software, hardware, archiving and data transfer, training, and the future (Grant 1986, pp. 21–31). This approach, of carrying out national surveys and making recommendations, has been followed elsewhere, for example in Spain (cf. Martinez & Lopez 1989).

Unlike the USA, where archaeologists properly trained in quantitative methods are seemingly legion, the use of statistical and simulation approaches (Hodder 1978; Sabloff 1981), with some exceptions, is largely confined to the academic wing in the UK. The exceptions generally stem from central bodies who have influence over other groups in the regions. For instance the Ancient Monuments Laboratory (AML) produced a package for researchers specializing in faunal remains to record and process ecofacts (Jones, Locker, Coy & Maltby 1981) and the Central Excavation Unit (CEU) supplied several sites and monuments recording packages (eg. Booth 1988) which, previously, other units had to adopt in order to receive funding from London. Of course, individual groups have bought various other computer packages on an *ad hoc* basis, rather than as part of the rationale of some long term plan. It must be said, however, that the whole situation is changing rapidly because of the greater professionalism adopted in the deployment of computer-based applications.

Academics have seemed more prepared than other archaeological groups in the UK to regard computer-based research as a sub-discipline to be taken seriously, and even to be fostered through the holding of special meetings (eg. CAA), the establishment of newsletters (eg. *Archaeological Computing Newsletter*), electronic conferencing (eg. *Archaeological Information Exchange*), and special awards and grants, through bodies such as the British Academy and the Science and Engineering Research Council's (SERC) Science-Based Archaeology Committee (SBAC) (Smith & Harris 1990). Certainly, there are a growing number of departments where individuals can learn or develop computer-based archaeological skills. Besides the Research Centre for Computer Archaeology in Staffordshire Polytechnic, which has been in existence for many years, there are now a number of universities, such as those of Southampton and Leicester, which offer postgraduate diplomas and Masters degrees in computer-science, or IT, and archaeology.

Nevertheless, there remains a time-lag between the inception and presentation of new computerized methods, and their permeation into daily practice, this time lag being in fact quite variable. In recent years, database structures, seriation and stratigraphic sorting algorithms have been advocated and adopted in British field units at an impressive rate: CAD/CADD (eg. Huggett 1990) and GIS (see Lock & Harris Ch. 9, this volume) applications are currently in vogue. Here too, the situation is changing rapidly and many more people working in British archaeology will, in the future, receive significant education in the use and abuse of IT. Today, great emphasis is also laid on training in the use of computers at secondary and higher education level (eg. Bennet 1986; Martlew 1987; Greene 1988; Richards 1988; Dobson 1990), as well as on training for professionals. However, we must wait one or two student generations for these people to rise through the ranks before we will be able to measure any real impact on the approaches adopted in Britain's various archaeological organizations.

The British situation is just a subset of the global picture, and has its own peculiarities. A more panoramic view of the subject reveals many interesting facets.

How is information technology used in archaeology?

Over the decades archaeologists have become increasingly sophisticated users of IT. Gaines (1987) characterized the 1960s as the age of 'exploration' of computing archaeology, the 1970s as the age of 'implementation', and the 1980s as the age of 'exploitation'. The 1990s are destined to be characterized as the age of 'information', as the majority of research has now moved largely to the newer enabling technologies of databases, graphics and artificial intelligence, and away from analysis to field recording and publication.

Even a classification scheme as general as this is procrustean, since, even within the most industrialized and economically powerful nations of the world, there are teams, or subject areas, within which IT is still ignored, or is only just beginning to be explored, implemented or exploited. However, although IT is not an integral part of archaeology around the globe, it is nevertheless all-pervasive to the extent that it is now possible to document digital skeuomorphs of virtually every major archaeological procedure.

IT should be attractive to archaeology because it mediates between a multitude of activities associated with archaeology and archaeologists. Hence, we can speak of electronic field recording (eg. Korsmo, Plog, Upham & Cordell 1985; Alvey 1990; Fletcher & Spicer Ch. 10, this volume; Oikawa Ch. 6, this volume), electronic laboratories, electronic offices, electronic archives (eg. Martlew Ch. 21, this volume), electronic mail (Anon. 1991), electronic conferencing (eg. Rahtz 1986; Rahtz 1989), electronic networks (eg. Jakobs & Kleefeld 1991), electronic books (eg. Makkuni Ch. 20, this volume; Rahtz, Hall & Allen Ch. 22, this volume), electronic class rooms (eg. Richards 1988), electronic or, to use the current terminology, 'smart' museums (eg. Roberts 1990; Makkuni Ch. 20, this volume), and even electronic sites (eg. Ruggio 1991; Molyneaux Ch. 19, this volume; Reilly Ch. 12, this volume). The technology used here is not necessarily passive. Computers will trap errors during data entry, they will enhance poorly defined aerial photographs and other remotely-sensed 'images' (eg. Haigh 1989; Scollar, Tabbagh, Hesse & Herzog 1990), they will even argue with users' reasoning—admittedly within tightly defined domains of discourse (see Stutt & Shennan Ch. 18, this volume). Moreover, the technology is becoming both more portable and robust; notebook-and palm-sized computers now possess enormous processing power. They are even cryogenically tested. Modern hardware has been shown to withstand the environmental and climatic extremes of deserts (eg. Knapp 1986), tropical rain-forests (eg. Roosevelt 1987), and arctic wastes, and thus can be safely carried to the most inhospitable and inaccessible parts of the planet. Alternatively, it may be more effective in some instances to study certain archaeological formations remotely. For instance, digital satellite images enable the tracing of lost sections of the Great Wall of China (Guo forthcoming), while advanced graphic systems have facilitated the visualization of underwater assemblages in the Great Lakes of Canada (Stewart 1991).

An important channel by which IT has been introduced into archaeology has been via field-data collection. The development of microchip technology precipitated the use of lap-top computers and data-loggers in archaeology, although caution against their uncritical acceptance into the professional's tool-kit has been expressed (cf. O'Brien & Redman 1985; Reilly 1985; Stephen 1987). Huggett's (1990, p. 3) description of the IT base of the Deansway Project, a recent excavation in the city of Worcester, UK, typifies the situation in many modern excavation projects:

> Computers have had an impact on almost every area within the Project —illustration, display work, finds processing, site planning, training, and public relations. In each case, the aim is to enhance the work being done in terms of speed and quality. The computers are not used by computer experts as replacements for specialists, but the specialists themselves are training in the use of the equipment.

Many of the traditional tools of the field archaeologist now have a digital counterpart. The humble pencil-and-pad combination is increasingly being supplanted by the hand-held machine containing pre-programmed *pro-forma* systems, often supplemented by an array of error-trapping procedures. In the near future, we can expect even greater latitude in our input methods, including reverting back to cursive handwriting using the latest generation of user interfaces currently on the market (cf. Carr 1991). Advances in speech recognition make it likely that voice-entry and -output will be a generally available option in the foreseeable future. The shift away from the keyboard to other forms of human-computer interfaces is a step towards the greater democratization of knowledge, since people with non-existent or poor keyboard skills will not be disadvantaged. Those who wish to transfer knowledge by speech, rather than writing, are an obvious group who will benefit from these developments. For many groups—both individuals with no formal education as well as non-literate societies—communication is advantaged by not being formalized in writing (cf. Layton 1989) and, for them, real benefits are to be gained.

Progressing hand-in-hand with input methods are output methods. The development of the ingenious output devices which currently astound or amuse, but which undoubtedly will be regarded as unremarkable by the year 2000, already enable such powerful human-computer interfaces that a human can be fooled into believing it is entering into the electronic world of a so-called 'Virtual reality'. Dynamic audio and visual feedback are already available for some applications and tactile feedback is possible. Undoubtedly, the computer simulation of an excavation in which the archaeological trainee is clad in data-gloves and data-helmets is on the horizon.

The most obvious advantage that computer technology has over human skills is in the context of mechanical, repetitive, but crucial, tasks such as seriating finds and unravelling the complexities of large stratigraphic sequences, operations which can be carried out faster and sustained over far longer periods by computers. The literature abounds with seriation techniques. The *Bonn Seriation Package* includes an extensive collection of relevant citations (Scollar 1989). The basic components of a Harris Matrix are habitually ordered by a large variety of stratigraphy processing programmes (eg. Ryan 1988; Rains 1989; Alvey 1990; Boast & Chapman 1991; Desachy & Djindjian 1991). This considerably enhanced facility for data verification sometimes has unexpected side-effects. An interesting discussion of the impact of computerization on site archives is provided by Mc Vicar & Stoddart (1986), who encountered worker hostility to over-efficient feedback.

Field drawings, such as the plan and section, are already in various stages of automation. Computer-aided design (CAD) and computer-aided drafting and design (CADD) packages are not only used for good quality 'reconstructions' (eg. Eiteljorg 1988; Durliat, Saint-Jean, Kuentz & Lorimy 1990; Huggett 1990; Templeton 1990). Schematic drawings, such as single context plans, may be digitized, using a digitizing tablet, and manipulated using these systems. Such plans can be electronically linked to other field data held in databases (Alvey 1989; Alvey 1990). Photographic records can also be digitized with digital cameras and handled in analogous ways (Stančič 1989). 'Drawings' can be augmented with other topographic data, such as those collected using electronic theodolites (eg. McPherron & Dibble 1987). These instruments are used most extensively and intensively for recording the three-dimensional locations of artefacts and ecofacts from excavations.

As the quantity and complexity of the data grows, increasingly sophisticated analysis and display methods can be called upon in investigations to help researchers explore and understand the form, structure and content of formations. The methods range from simple two-dimensional graphs and scattergram charts, through two-dimensional slices from three-dimensional assemblages (cf. Upham 1979), through three-dimensional interactive object-orientated vector graphics (eg. Colley, Todd & Campling 1988), to full three-dimensional interactive volume visualization systems, including volume renderers (eg. Stewart 1991). The technique adopted will be determined by the data, the available hardware and software, and the questions being asked. Visualization methods are very qualitative, but they have leap-frogged quantitative approaches which have until now been preoccupied with analyzing two-dimensional distributions. Koetje (Ch. 13, this volume) attempts to redress the balance with his innovative application of simulation to recover archaeologically meaningful levels in deposits.

Many devices used by field archaeologists (eg. proton magnetometers, resistivity meters, radar and other geophysical survey equipment, cameras, etc.) will output their data in digital format (eg. Hartman 1988; Anderson 1991). Whereas machine processing limitations have meant that survey data would be presented using established display techniques, such as two-dimensional contours or dot-density plots, several scalar data sets can now be displayed on wireframe or shaded polygonal-patched graphical models (eg. Arnold, Huggett, Reilly & Springham 1989; Reilly 1989; Voorrips & Ottaway 1990; Richards 1991; Fletcher & Spicer Ch. 10, this volume).

Not surprisingly, the processing of small finds which, in the past, involved considerations associated with cleaning up, preliminary conservation and data logging, has not escaped the drive towards greater efficiency. Artefacts can be registered in seconds using digitizing cameras or video cameras attached to frame grabbers, and can then be transferred to a bar-coded storage container (cf. Kampffmeyer 1986; Plog & Carlson 1989; Kann-Rasmussen & Larsen 1991). Three-dimensional digitizing devices such as the McDonald Douglas 3SPACE Tracker are used for registering the decoration on ceramics (Oikawa Ch. 6, this volume) and for recording the soil silhouettes of bodies buried at Sutton Hoo (Reilly & Richards 1988; Reilly, Richards & Walter 1988). Examples of integrated systems involving digital data collection out in the field, and on-site data-processing, have been described by Powlesland (1983, 1985), Plog & Carlson (1989), Weiss (1989) and Huggett (1990). The management of the resulting digital archive has been discussed, for example in relation to the French Archéodata project (Zarzosa & Arroyo-Bishop 1989). Other database management systems have been engineered specifically for archaeological purposes. The MINARK database program is employed not only in Australia Johnson 1984b), from whence it derives, but also in Europe (Ottaway, *pers. comm.*), East Africa (Sinclair, Meneses, Kokonya & Rakatoarisoa Ch. 2, this volume), North America (Johnson, *pers. comm.*) and the Middle East (Knapp 1986; Weiss 1989). Others prefer to work with 'off-the-shelf systems (eg. O'Neil 1984).

Of course, it is not only the field archaeologist who is harnessing the power of IT. The specialists working in laboratories now make full use of the speed, power and additional facilities that IT can bring to a laboratory environment. Measurement is fundamental to any laboratory. The variety of measurements is almost limitless, from grain size in thin sections, through to the thickness of fibres in textiles (Ryder 1986), to the three-dimensional form of decoration on ceramics (eg. Kampffmeyer 1986; Oikawa Ch. 6, this volume).

In all the applications mentioned, IT has proved to be a useful tool. However, it is really at the organizational level that its power and flexibility come to the fore. In large and complex institutions, like planning departments, large museums, universities, and so on, effective information systems are essential if the organization is to function. In this kind of environment IT does not simply supply a series of useful tools, it is sublimated into the infrastructure of the institution. Technological osmosis occurs in many ways, but particularly out of the administrative and management section, where day-to-day data management activities can be catered for by word-processors, spreadsheet and database programs, financial packages and the other systems which go to make up the electronic office—and the benefits are associated with efficiency and saving of costs. As large institutions adopt IT for administrative reasons their constituent departments will be obliged to conform to the new working conventions. These changes gradually diffuse into other activities. In many countries, archaeological organizations are part of larger groups (such as universities, local council planning departments etc.), and are under pressure to be cost-effective and forward-looking. Thus the transition to on-line working practices—all data, forms, letters, etc. being processed through computer terminals—results from non-archaeological reasons, in contrast to the introduction of IT for purely research purposes.

Other reasons for adopting IT have less predictable outcomes. For instance, individuals may be presented with 'hand-me-down' systems from other groups which are moving on to more powerful equipment, or are given equipment by colleagues. Sinclair, Kokonya, Meneses & Rakatoarisoa (Ch. 2, this volume) recount two interesting variations on this theme. The first concerned the donation of Osborne microcomputers to Madagascan archaeologists from north American colleagues, the second donation was from Cuba to Mozambique and was given for political reasons. In neither case was the choice of machine based on the local archaeological needs. In addition, it is important not to underestimate the powerful driving force of the individual ego—innovation is often produced by one dedicated person. Often, the public image of a project such as the All American Pipeline Project (Plog & Carlson 1989), or London's Department of Urban Archaeology (Boast & Tomlinson 1990), is considerably enhanced by its computing prowess, and this again can be an important factor when organizations are competing for public funds, or working in an enterprise culture of contract, or rescue, archaeology. The majority of archaeological work does not provide spectacular public displays like Akrotiri, Macchu Pichu, Great Zimbabwe, Stonehenge or Angkor Wat. Therefore, the archaeologist seeking attention must enliven the work in other ways, such as by making colourful computer reconstructions (eg. Wood & Chapman Ch. 11, this volume)—though this is not to imply that the reconstructions do not have considerable research value—creating interactive museum displays (Makkuni Ch. 20, this volume; Maytom & Torevell 1991), or even showing hardware in exotic places (Powlesland 1983). The attraction of media attention draws significant numbers of the general public into an interest in archaeology, as well as persuading potential sponsors that archaeology is a worthwhile area to invest in.

This shift towards IT-based storage and dissemination methods, and the changes they may entail, are partially due to the shortcomings in existing recording and dissemination methods. Archaeological publication is in a state of crisis in most countries and regions. At one end of the spectrum are those countries where large numbers of rescue or salvage excavations, carried out prior to development, are generating correspondingly large numbers of archives and reports. The problems which occur in this environment are not new and have been discussed at length in the Anglo-American orbit (eg. Gaines 1984; Hassall 1984; Oppenheim 1984; Powell 1984), and in relation to Japan (eg. Oikawa Ch. 6, this volume). Trifonov & Dolukhanov (Ch. 7, this volume) are concerned that the mass of information remaining unpublished in the USSR has resulted in foreign scholars developing totally fallacious explanations of the role and nature of Soviet archaeological material in the general context of interpretations of both European and Asian archaeology. Elsewhere, in central Europe, only a small percentage of reports ever get published, and the long delays even in the publication of these is a cause for concern (eg. Marciniak & Rczkowski Ch. 4, this volume). This story is echoed in east Africa (Karega-Munene Ch. 3, this volume; Sinclair, Kokonya, Meneses & Rakatoarisoa Ch. 2, this volume). The basic problem is that the number of reports being written is swamping existing publishing oulets.

Some people are convinced of the benefits of adopting IT for collecting, processing and disseminating data and interpretations, and regard the traditional methods of making archaeological data available—books, journals, microfiche—as cumbersome, inadequate and expensive. In Third World countries, expense means that availability is usually highly limited, involving only a very small number of libraries and academic centres (eg. Wandibba 1990).

The advent of desk-top publishing is having a profound effect on archaeological publishing. Desk-top publishing emerged in the mid-1980s, when relatively inexpensive microcomputers with comparatively large memories appeared on the market together with moderately high resolution laser printers (Girdwood 1988). Today, many reports are generated and updated frequently in the field. In the specific case of contract, or rescue, archaeology, the use of desk-top publishing to produce attractive proposals and reports quickly is crucial to success in the ‘bidding’ process. Many small units now regularly produce high-quality, camera-ready documents for private publication. As the overheads of desk-top publishing decrease, this trend will grow. As more archaeologists are in a position to prepare their own camera-ready documentation, some of the bottlenecks of type-setting at publishing houses can be by-passed. However, it is recognized that the established, almost ritualized, process of academic publication in learned journals and books offers many advantages to the would-be author (eg. credibility in the eyes of the relevant peer group through the refereeing process, and wider exposure through organized publishing outlets); nevertheless, the need to publish quickly, and at several levels (eg. glossy books for lay visitors to a site or museum, technical reports for local specialists, and general reports for professional consumption), is often a reason for the introduction of IT. In fact, a growing number of archaeologists are circumventing the larger publishers, who have little interest in publications with print-runs which seldom exceed a few hundred units, by printing their own reports and distributing them themselves through specialist archaeological outlets. A side-effect of this shift from controlled to uncontrolled publishing has been to undermine or challenge the traditional power of the academic editor.

It can be argued that the proliferation of publications fosters the intellectual vitality of the subject by allowing far more views and positions to be expressed. Conversely, as many more publications are not being subject to extensive peer review before publication, readers, faced with growing mountains of archaeological literature competing for attention, have little to help them determine the standard of work embodied in them. Also, the process of discovering those works of most relevance to a particular reader is made more difficult by simple weight of numbers. This is not a new problem. There is a necessary tendency in most fields to accept derived data. Instead of checking raw data, the personal standing of the author of the interpretation is used to validate it—authority X said ‘…’, so it must be true. Trigger (1989, p. 16) has observed that:

> …while archaeological data are being collected constantly, the results are not necessarily as cumulative as many archaeologists believe. Indeed, archaeologists often seem to build more on what their predecessors concluded about the past rather than on the evidence on which these conclusions were based.

The availability of electronically vendered data could change this situation drastically by allowing the free flow of primary data.

In the future, the position of ‘editor’ may become untenable if, as some believe, hardcopy archaeological publications are superseded by electronic books, and as the addition of motion video and interactive three-dimensional graphics capability introduce still more, and largely unquantified, possibilities. But signs of change are indeed apparent. The publishing industry is gradually adapting itself to a more fluid form of dissemination in the form of ‘soft books’, both for popular titles and for academic publication. Already, some of the larger publishing houses have released compact disks (using CD-ROM technology, discussed *inter alia* in Lock & Dallas (1990)). Rahtz, Hall & Allen (Ch. 22, this volume), Martlew (Ch. 21, this volume), Makkuni (Ch. 20, this volume) and Lock & Dallas (1990) all discuss variations of the electronic book. Understandably, concern has already been expressed that moves in this direction might widen the rift between archaeology in the west and elsewhere in the world. Bagnall (1989) believes that it is possible to ameliorate the divisive side effects of this disparity by ensuring that the electronic document is not the sole medium of dissemination. He suggests that those with the wherewithal to promulgate sophisticated IT-based information (eg. compact-disk books) should ensure that hardcopy versions of the books are also available, at least in the short term, so as not to exclude participation by anyone. While the sentiment behind this proposal is noble, it is unlikely to be practicable, given that such hardcopy versions are likely to remain limited editions (in publishing terms), and could thus only be produced at prohibitively high cost. In any case, this view denies what is probably the most important reason for adopting the new media, namely the freedom to break away from some of the constraints of representation and presentation imposed by traditional printed matter (Hodder 1989). One immediate advantage is the ability to include large numbers of pictorial data, particularly non-schematic images such as photographs, because it is cheap to do so (see Martlew Ch. 21, this volume). Once again, communication which does not rely on the written word is enabled. Makkuni (Ch. 20, this volume) shows one situation in which this enriches the process of discourse in a museum environment.

Conservatives would argue that there is a massive publishing industry and that most people are able to work quite satisfactorily through the medium of the printed page in which so much work has already been invested. A compromise would be to start out from a position which attempts to maintain the basic structure and form of a book, but would be much larger in terms of information content. Rahtz, Hall & Allen (Ch. 22, this volume) argue that there is a good case for this practice since people are well-used to looking at books and understand their tacit structure implied through the use of white space (eg. indents, blank lines). Standard Generalized Markup Language (SGML) conventions would enable the publisher to reproduce the style of a book, including headings, blank lines, indents, emboldening and other changes of font, to denote the different functions of sections (eg. quotes). The same SGML notation can be used to generate different styles, but also lends itself to exploitation as a navigational tool in a hypertext document, such as those envisaged in the projects mentioned in Chapter 22.

However, it can be argued that it does not make sense to publish a two-dimensional figure to illustrate a point about a three-dimensional subject when the full three-dimensional data set can be made available. The insights stimulated by the wide range of interaction and visualization methods mentioned above (and discussed in greater detail in Fletcher & Spicer Ch. 10, this volume) provide dramatic examples of this. The practice of referring to illustrations could be abandoned, and reference to figures from within a text could be replaced with a command defining what data sets and viewing parameters should be loaded. The point-of-view presented could be accepted or rejected, but the crucial difference is that the reader has the freedom to change position and modify the questions being asked.

It is possible to contend, with some justification, that the three-dimensional data archaeologists record are reduced to two dimensions in order to make them more tractable, even though modern data visualization techniques now enable such three-dimensional data to be examined dynamically in real-time from multiple viewpoints, and manipulated in ways which can reveal features hitherto unobserved. In short, archaeologists are now in a position to *explore* their recorded data more fully than ever before. The probable impact of such developments on primary archaeological recording are now receiving attention (cf. Fletcher & Spicer Ch. 10, this volume; Molyneaux Ch. 19, this volume; Reilly Ch. 12, this volume).

The material formations 'observed' and 'recorded' by archaeologists are not directly perceived phenomena but classes of intellectual constructs. Concepts like 'context' and 'association' are used so frequently that there is a tendency to forget their theoretical nature. The object of analysis and policy, as represented in archaeological records, is therefore theory-laden as a result of processes of definition, convention and abstraction. The use of tools (both material and conceptual) is a crucial factor since:

> their influence on the direction of work done is important and frequently decisive. New tools make possible the production of entirely new sorts of data and information. (Ravetz 1971, pp. 89–90)

In relation to this, Polanyi (1958) has drawn attention to the phenomenon whereby people skilled in the use of a certain tool or methodology are unable to specify in detail the skills they employ in accomplishing a task. Having perfected a new skill many procedural elements become tacit, and are undefined when the 'skill' is transmitted to another person, who may learn the bulk of these skills by example (Polanyi 1958, p. 53). However, while developing a new tool or methodology it is much more likely that a greater proportion of the tenets underlying the approach to the task will be made explicit. Hence, the introduction of a completely novel tool provides an ideal opportunity for some of the theoretical ground rules implicit in a subject to surface into the consciousness of the discipline, and be re-examined and assessed.

Archaeologists cannot assume that they simply transfer fossilized elements of past activity into a record where they can be further examined. Instead, elements are selected and then remodelled to suit the structure of the medium of storage. For instance, in Australia, Presland (1984, p. 7) states that 'implementation of a computer based register has allowed, and in some cases necessitated, a number of changes in the types of data recorded from archaeological sites'. The switch to computer storage implies that changes should therefore also be made in remodelling techniques. As a result, the composition and form of the finished record may be significantly altered. If so, this may warrant a reconceptualization of the definable data categories, entailing a reassessment of their epistemological status. Archaeologists who wish to adopt computers must be aware of these problems.

The history of archaeology records the activities of archaeologists and their individual interests. Naturally, the data sets they produced were to some extent shaped by the nature of the questions they asked and the methods they employed to answer them. Thus, the information they procured is value-shaped and theory-laden. Archaeologists have consequently had to develop many tools with which to extract information from the material of others. In making comparisons or syntheses, archaeologists may utilize, criticize, ignore, or simply forget the work of colleagues, both past and present. As methodology and theory have developed, large areas of archaeological knowledge have been discarded. Stockpiling data derived from a multitude of methodological and theoretical approaches in one or more large information systems could change all this. This is particularly true of systems that are described as intelligent. Increases in both data storage capacity and data handling ability advance the growth of a new cycle of development: the life of data collected in one intellectual milieu is being artificially extended and preserved. To deal with this, archaeologists will have to devise ways of extracting and accounting for alien epistemologies embedded in their databases. This restructuring of the nature of archaeological enquiry may not be wholly undesirable, but it must always be borne in mind that computers are not the sort of simple palliative that many appear to believe them to be.

It has been claimed that the use of artificial intelligence (AI) techniques and formalizing techniques enhance archaeological reasoning processes by forcing practitioners to think differently about the underlying nature of their assertions, interpretations and inferences (eg. Huggett 1985; Doran 1987; Voorrips 1987; Gallay 1989). Doran (1987, p. 87) has gone so far as to suggest that:

> within archaeological training there is needed some shift of attention away from relatively sophisticated multivariate statistical methods towards computational and AI techniques. No archaeological student should be left unaware of the potentialities of information science in general.

The formalization and insights that Dallas (Ch. 16, this volume) brings to the study of the classical Attic stele should be set against the backdrop of a discipline in which the subject matter has been the focus of continuous and intensive research by generations of scholars. The profound new observations he is able to make testify to the fundamental relevance of such formalization methods to help archaeologists identify and explore productive new areas of enquiry. Biskowski (Ch. 15, this volume) shows another potentially fertile area of the application of formalization methods. He harnesses the power of AI techniques to enable him to build controllable simulation models. One attraction of this approach is that it becomes less difficult to monitor and understand these notoriously complex systems. Yet, despite the enthusiastic assessment of worth and relevance of AI techniques by the relatively few archaeologists who work in this field, archaeologists remain largely blind to the urgings of Gardin (eg. Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988; Gardin 1990; Gardin 1991) to formalize their language and to provide environments in which archaeological reasoning can be embedded and analyzed, allowing formalizations of historical processes which can be discussed logically. Biskowski (Ch. 15, this volume), Dallas (Ch. 16, this volume), Lagrange & Vitali (Ch. 17, this volume) and Stutt & Shennan (Ch. 18, this volume) describe four very distinct projects which demonstrate, in a number of different ways, that the use of formal methods can allow the structure and contradictions or flaws in archaeological reasoning to be made explicit, and thereby may stimulate new insights.

Generally speaking, the few expert systems that do exist in archaeology are restricted to very well-defined problems such as the appropriate conservation treatment for damaged frescoes at Pompeii (Quattrin 1990) or the dating of well-known artefact typologies (eg. Ozawa 1989). Lagrange & Vitali (Ch. 17, this volume) describe an expert system which was developed to help archaeologists in the interpretation of results from technical analyses within the framework of archaeometric provenance studies. Their experiences lead them to conclude that the system should be made available to the archaeological community both as a research and a tutoring aid.

That these systems have been exploited successfully, to distil a body of factual information and a series of procedural rules (of the kind: if…then…) to adequately represent specific domains of archaeological discourse, lends some credence to Gardin's (1990) claim that archaeological publications could be drastically reduced in size if they were transformed into so-called logicist schematizations, which are indisputably similar in structure to rule-based expert systems. Stutt & Shennan (Ch. 18, this volume) believe that they can improve on what they describe as the prescriptive and positivistic arguments embodied in the logicist approach by embodying some of the capacities of archaeologists as 'arguers'. In so doing, they suggest that they can 'expose' and, thereby, create the opportunity to improve the nature of arguments, while, at the same time, capturing a flavour of complex academic dialogue and argument.

Already, the analytical possibilities offered by the new methods may have potential repercussions for both data collection during the course of archaeological fieldwork and, consequently, through their incorporation into interpretations of past human society. However, before introducing a new technology, potential users should consider in detail the implications of such a move. The introspective questions posed by the early pioneers and advocates of IT in north America and western

Europe are still pertinent today. In assessing the potential impact of the introduction of a new technology a number of questions should be asked: Why use a computer? Should it be used? Most important of all, is anything going to be accomplished by the use of computers which cannot be done just as well by some other means? (see Chenall 1968, p. 17). There is a danger that computer technology may be embraced for the wrong reasons, with undesirable side-effects. For instance, craftsmen may be pressurized into taking on board a tool solely in order not to lose their jobs. Some may embrace technology, such as expert systems, to avoid the rigour of reflection (the same criticism could be made of the use of statistics!). Administrators may view IT as a means of *control*—centralized data is easier to withhold (the same criticism could be made of libraries—would we castigate medieval European monasteries for hoarding knowledge or praise them for preserving the classical works?). Some groups might adopt new IT for appearances sake. Electronic communication, for instance, effectively provides membership of a privileged club.

The growing availability of computer networks in many countries (eg. Jakobs & Kleefeld 1991), as a result of economic and other needs, may mean that such communication will be increasingly available to archaeologists in all parts of the world. If archaeological data dissemination, and archaeological dialogue more generally, is intended to take place over such networks, is it simply naive to hope for what might be termed the democratization of archaeological knowledge on a global basis? Alternatively, are we on the brink of a period in which transaction processing, usually associated with the banking and merchandizing business worlds, will be necessary to exchange data, and does this herald the rise of specialist brokers of archaeological data? Electronic mail and remote access to sites and monuments or bibliographic databases, for example, are possible in many western countries. As these networks gradually link up and extend, two major issues emerge: one is the allocation of costs in cooperative-and distributed-processing; the other concerns questions of ownership and copyright. Much research effort has already been directed to the more effective use of existing processing capability. Partitioning the processing tasks over networks of processors (which are often from different manufacturers) is one alternative being explored. Accounting for who uses what processing power where, will force the contributing partners closer together administratively. The debate over copyright focuses on those archaeological organizations who wish to make charges for the use of ‘their’ data; claims to primacy of ‘ownership’ of archaeological material is a potential growth-point of friction.

Networking produces enormous strains within the structures and relationships of archaeological institutions and organizations as they become more integrated into local, national and international electronic networks. There is the perennial problem of standardization. The word has many meanings, depending on the context in which it is used; Richards (1985), for example, discusses problems in establishing and maintaining standards, making an important distinction between hardware and software standards and data standards; Wilcock (Ch. 8, this volume) considers the further issue of data exchange standards. As the IT world gradually moves to so-called ‘open systems’ —where, in theory, users are free to move or share software systems over different manufacturers’ hardware platforms—it is inevitable that bigger organizations will impose their own standards. In a scenario where the infrastructure of several, perhaps rival, archaeological institutions is highly dependent on IT, and those institutions can be connected electronically via networks in a seamless fashion, using open systems and distributed-processing, such organizations will inevitably become inextricably bound together, possibly even exaggerating political tensions. Already this can be seen to be happening. The Museum Documentation Association’s MDS (museum documentation system) is now a *de facto* standard for museums around many parts of the world. In Europe, electronic networks mean that heritage registration systems such as ArchéoDATA (Arroyo-Bishop 1991) operate in several adjacent countries. The recent re-unification of Germany is likely to throw up many problems in the process of literally amalgamating museums and museum collections which were split up as a result of the partition of Germany after the Second World War. The situation is poignantly symbolized in Berlin’s Egyptian Museum, where the bust of Queen Nefretiti, in what was West Berlin, was separated from her husband, King Akhenaten, and their daughters, who were held in East Berlin. The highly automated collections of what was formerly the west have to adapt to the circumstances of the collections which were located in the east. Ultimately, a country’s various archaeological institutions, with their different briefs, will coalesce into one meta-organization: university departments, archaeological field units, museums, planning departments and other governmental departments will be inseparable. The notion of a national strategy has in some instances proved elusive—because of opposition to the idea of a uniform policy, on the grounds that it would stifle, if not wholly destroy, the diversity of views and approaches which are embodied in the humanistic tradition of the subject (eg. Reilly 1985; Richards 1985), or because it has been impossible to identify who should take responsibility for the strategy (eg. Cooper 1985). However, this may now become a realistic objective.

Regional patterns?

Many archaeologists, particularly in the west, eulogize the merits of IT. However, if the eponymous technology of the new age of archaeology is obviously so suitable for the multifaceted activities of the discipline, why is it that so much of the archaeological world is so thinly populated with the technological wonders we have outlined above? It is clear that the systems and technology described above are not available or used everywhere on an equal basis. Yet the actual details of where and how IT is being used around the world generally is difficult to gather. There is no central electronic conferencing medium with which to monitor global trends.

How and why has IT become implanted in archaeology? Histories of the subject (eg. Wilcock 1989b; Djindjian 1991) imply that the prime mover was research interest. Overviews of specific countries have been offered for Australia (Johnson 1984a), China (Wilcock 1989a; Guo forthcoming), Denmark (Andresen 1988; Madsen 1988), France, Portugal, and Andorra (Arroyo-Bishop & Zarzosa 1989; Djindjian 1991), Hungary (Biro 1990; Suhadja Ch. 5, this volume), Poland (Marciniak & Rzkowski Ch. 4, this volume), Spain (Arroyo-Bishop & Zarzosa 1989; Martinez & Lopez 1989), and the USSR (Trifonov & Dolukhanov Ch. 7, this volume). In the USA, with its long and intimate relationship with computer-based methods in archaeology, assessments of their impact and statements about direction have been made as each decade passes (eg. Chenall 1968; Whallon 1972; Gaines 1984). In north America and western Europe the adoption of IT soon after its inception came to be regarded as a positive development. Whallon (1972, p. 30) believes that ‘computer processing has largely influenced analysis of archaeological data by permitting more complex manipulations than could ever be achieved in any other way.’ Additionally, several authorities stress ‘the generally beneficial effect of exposure to…computers’ (Doran & Hodson 1975, p. 346). Gaines (1984, p. 72) reinforces this view when commenting that the use of information systems by American archaeologists during the 1980s has ‘increased the need for clarity and precision in terms of data and procedures.’ Both Gaines (1984) and Whallon (1972) agree that ideas and approaches derived from the world of computers are forcing archaeologists into new ways of thinking. Moreover, it is clear that this experience has contributed to the production of much valuable work, especially through analytical studies using computers.

Of course, these might not be sufficient reasons to justify investing new, or diverting existing, resources to enable new technology and methods to be adopted. The role and relevance of archaeology in a society, or a section of society, may be such that any move towards adopting high-technology solutions is seen to be a public display of sympathy for foreign values and ideology, which may be regarded with alarm or suspicion. Political, ethnic and cultural attitudes to archaeology and heritage are clearly important to the discipline’s development in any particular area (cf. Layton 1989; Cleere 1989; Gathercole & Lowenthal 1990). How archaeology is taught within an education system of a country is one indicator of how the subject is regarded, but the importance of archaeology in any national curriculum is not only in terms of, for instance, explaining, justifying or maintaining political, cultural or ethnic identity (cf. Stone & MacKenzie 1990). Archaeology can be taught as a vocation or as part of a general education. In the non-vocational realm, archaeology is considered a useful vehicle for transmitting and developing skills related to teamwork, management and data handling at a variety of levels. In schools it must compete with other— so-called mainstream—disciplines such as history and geography. Teachers stress the suitability of archaeology as a flexible medium for applying and developing basic skills in, for instance, computing. Some find this ironic as professionals repeatedly try to disassociate themselves from projects which smack of scientism looking for an application. In the face of competition which may take away an educated support group and flow of potential archaeologists, some archaeological teachers are actively promoting the discipline. Martlew (1987, p. 1), for instance, has been very forthright in his opinions concerning the promotion of archaeology as good general education:

> Archaeology is facing greater threats now in both schools and universities than at any other time since it appeared on the timetable…. Educational policy makers need to be presented with a new image of the subject, and, however cynical it may sound to suggest it, the introduction of computers into schools has provided a convenient bandwagon on which a new image might ride. Archaeology can no longer be represented by the Indiana Jones Guide to Museum Collecting. It must be shown to be a subject which uses scientific techniques in its study of human societies, which has a strong practical bias, and—for the educationalists—which can be used to develop skills which are relevant in many other areas of the curriculum.

Many might assume such considerations to be of an exclusively ‘western’ kind so that Oikawa’s (Ch. 6, this volume) sketch of current Japanese practice is particularly illuminating. Almost all of his discussion concerns visualisation and reconstruction, but it is clear that most major institutions are turning to IT to help them meet the challenge of processing the colossal volumes of data being generated by Japanese archaeologists. Japan is similar to other capitalist countries in the range of application areas in which IT is found. Wilcock (1989a) gives a visiting western scientist’s impression of the depth and scope of computer applications in China; his perception was that of a country still in its infancy as far as day-to-day archaeological applications of IT are concerned. Guo (forthcoming) provides an interesting contrast by a Chinese national. He shows that archaeology in China is prepared, and able, to utilize IT. The major applications of this technology are, however, largely confined to projects associated with prestigious monuments, such as the Great Wall and the terracotta armies, and national museums. The fact that the Chinese departments of archaeology do invite foreign specialists to assist in technology transfer into the ranks of Chinese archaeology indicates their commitment to fully assess and make use of relevant techniques. In the southern hemisphere, Australia (Johnson 1984a) has been inculcated with IT-based applications in much the same way as north America, if not to the same extent. Australia is distinguished by having developed a uniquely adapted database management system for archaeological purposes (Johnson 1984b).

In other countries and situations, there are markedly different attitudes to, and uses of, IT. The descriptions of the situation in Poland (Marciniak & Rączkowski Ch. 4, this volume) and the Soviet Union (Trifonov & Dolukhanov Ch. 7, this volume) are also applicable to Czechoslovakia. In all these countries only a few, old-fashioned computers are used in the context of archaeological enquiry. The consequent restrictions to access mean that it is impossible for most people to be exposed to, and become properly trained in, computer applications. In Czechoslovakia, at least, the problem is alleviated through a series of collaborations with mathematicians and computer scientists (Bures, *pers. comm.*). The selection and direction of IT-based projects are thus more sharply focused in a few particular directions, unlike in western, capitalist countries which have been able to afford to explore a greater variety of technologies and application areas. In all three of these countries, discussion of the applications and impact of IT is not through the traditional mechanism of the printed page, but through inter-personal contact in the workplace or at conferences and other organized meetings. It is also interesting to note that in the Soviet Union, the technology is not really regarded as a potentially active—modifying—medium, but as a benign facilitating medium. According to Trifonov & Dolukhanov (Ch. 7, this volume) the main concern is to project modern Soviet archaeological theories, and the evidence which underlies them, to a wider audience through the use of IT. The picture painted of Hungarian IT-based archaeology by Suhadja (Ch. 5, this volume) is significantly different from other central European countries, for here there has been a long history of quantitative analysis, but only recently has work on databases really gained a firm foothold.

The situation in the UK has been described above. Within Europe there is a fair amount of technology transfer although even there, local circumstances, especially the organizational structure of the various national archaeologies, profoundly affect the manner in which IT is adopted. Denmark fits firmly into the mainstream of European archaeological practice but given their relatively small and close-knit archaeological community, the Danes have been able to establish national standards and strategies (eg. Andresen 1988; Madsen 1988; Kann-Rasmussen & Larsen 1991). Such national standards have also been established in the totally centralized archaeological service of the tiny state of Andorra (Arroyo-Bishop & Zarzosa 1989). This has not proved possible in Spain and Portugal, however, which have a number of autonomous regional archaeological organizations. France is the epicentre of the logicist school (Gardin 1988) which has made extensive use of expert systems in recent years, giving a distinctive flavour to the French archaeology scene. France, Holland, Belgium, Germany and Italy have produced a series of scholars contributing to the application of novel IT as well as formal methods. The current situation in Italy is of particular interest as government support has enabled the establishment of a number of major IT-based showcase excavations and museums (eg. Bruschini & Politi 1990), and the Italians look to the future of the sub-discipline with particular confidence (cf. Guio 1991).

A quite different perspective is gained from the review of the current situation in eastern Africa, where quantitative methods and IT are being introduced into what are largely virgin territories as far as familiarity with the uses and potentials of this sort of archaeological thought is concerned (Sinclair, Kokonya, Meneses & Rakatoarisoa Ch. 2, this volume). The excitement and success of this project also, of course, creates problems of training, cost and information dissemination which are discussed in a specifically Kenyan context by Karega-Munene (Ch. 3, this volume). It is interesting to note that in regions of the world such as Africa, where the population density is lower than in industrialized parts of Europe, and where excavating small sites is not effective, the focus is on regional surveys, accompanied by prospection techniques and the setting up of computerized databases. In the Sahara large scale surveys are achieved through the use of landsat images (Ammerman 1989).

The above picture is one of great variety. The common themes to have emerged concern data management, data exploration, and data dissemination.

Democratization of archaeological knowledge?

We can distinguish at least three distinct aspects of intellectual material associated with IT: ‘data’, ‘techniques’ and ‘knowledge’. Data are the basic building blocks of practical archaeologists, even when they form high-level structures like the distribution of roman coins in Europe, the finds of sassanian pottery in east Africa and the evidence for the husbanding of pigs in Malaysia. Techniques are essentially the methods and tools used. These include laser theodolites or resistivity meters, the construction of databases, the use of multivariate statistics, and even the schematizations of logicists. Techniques applied to data (ie. capture, manipulation, storage and retrieval) provide more useful information (eg. a set of spot heights is converted into a terrain model or contour plan). These pieces of information are generally of little use on their own. Selected pieces of information can be used as evidence to support particular interpretations and conclusions which must be disseminated and accepted before they can be said to be part of the knowledge base of a subject (see Ravetz 1971; Majone 1980). The relationship between these entities is not straightforward, being neither constant in time nor place.

Data, techniques, and knowledge are inextricably intertwined in an intellectual milieu, but not all of them are equally amenable to sharing. The central issue is one of ‘access’, which itself produces a bewildering array of problems: how may the archaeologist from Colombia, for example, gain access to the minutiae of the bone records from the iron age of the Thames valley in the UK? Lurking in the background of the emotive issue of access is the ever-present spectre of the intellectual imperialist, who is capable of acquiring material or restricting access to it for ‘personal’, ‘political’ or ‘religious’ reasons (cf. the widely publicized dispute over the delay in publication of the ‘Dead Sea Scrolls’—Hebrew texts which may throw new light on early Christianity). The adage ‘knowledge is power’ may be true, but access to techniques is a prerequisite. Discussion of ‘knowledge’ is not only a matter of knowing the whereabouts, or even the existence, of information but, crucially, involves also ‘access’ to such data and techniques. When access to computerized data is granted freely, practice may change. Access by itself does not ensure that a technique will attract the interest of potential users and thus lead on to testing, acceptance and eventually refinement. Potential users have to be convinced of some tangible benefit before adopting a new technique or tool. The problem is particularly acute in leading-edge technologies which may be regarded as, at best, fringe activities or, at worst, intellectual abominations spawned by outsiders—those not perceived as mainline archaeologists. In effect, a tool or technique must be *demonstrated* to the archaeological community. Published accounts of new methods or tools cannot convey the experience of actual use of the new technology, and the situation is not assisted by claims and counter claims about the function or genre, which may not be fully understood by the potential archaeological users. For example, when discussing ideas concerning the potential impact of ‘virtual reality’ methods, most people have no idea what the medium actually represents (see Molyneaux Ch. 19, this volume, for a discussion of this problem).

An obvious solution is to provide training. Presented with real, working, application (with actual data) it becomes possible to explore scenarios which parallel local conditions. The view from one African country (Madagascar) is succinctly put by Rakatoarisoa (1990):

> IT in archaeology has not been developed in a uniform way. Some countries are improving their systems while others are only at a starting point. There is, however, no need for the advanced countries to abandon their colleagues who are behind; the differences are not yet so great as to preclude any kind of dialogue. But we must be aware of the danger that the technological differences might become so considerable that it will become impossible to understand each other. There are probably several ways of avoiding rupture, but we suggest some simple actions: making the magazine and periodical subscriptions easier, providing access to the big databases, facing the development of training materials which can be adapted to the needs of each participant.
>
> It would undoubtedly be very useful to establish quickly a true relationship of collaboration and mutual trust between laboratories in different countries. This trust is fundamental to progress, and will increase credibility in the eyes of financing institutions, most of whom still doubt the capacities of researchers from the Third World to use computers. There has even been a case where a mission of experts was sent to evaluate whether Madagascar was capable of using such equipment and if there really was the need for it. Of course, after having sent this mission there was no money left to finance the acquisition of the actual equipment.

The enormously increased ease of transmission of information, which is developing with breath-taking rapidity in the First World, may leave the Third World even further behind. Alternatively, and more optimistically, the same technology may ‘bridge’ the information gap. A partial solution might be to explore synthetic data sets. Modern interactive, multimedia software is adaptable to modern training applications. It should be stressed that archaeologists taking advantage of such systems are implicitly training the next generation of archaeologists to regard digital data as ‘natural’. (The simulation of excavations using multimedia methods is discussed in Molyneaux (Ch. 19, this volume) while Makkuni (Ch. 20, this volume) describes a different kind of interaction with multimedia machines by museum-goers. Other multimedia experiments are reported and discussed in, for example, Grace & Orton (1989), Martlew (1989, 1990) & Ruggio (1991).) An immediate issue is whether to rely on locally generated simulations or to call upon the synthetic creations from outside. For archaeologists, having a machine to do local data processing is one thing. Having a machine which allows electronic access to large data archives, using networks or portable media (tapes, disks etc.), is quite another. The solitary machine is an invaluable aid to the archaeologist, assisting in the analytical tasks which take up most of a working day; interpretations and data are generated for dissemination in the usual way. But the addition of networks brings with it a new complexity and immediacy which can have wide-ranging implications: data can be shared with other workers in real time; conversely, they can also be kept private. In noting, once again, a relationship between knowledge and power, it is also salutary to recall the corollary that power necessarily involves *exclusion*. When knowledge is universal, it can be claimed that it has no power.

There are already substantial problems in knowledge sharing between archaeologists; the cost of books in many Third World countries is a straightforward example, excluding archaeologists from the work of their colleagues; also, the cost of attending conferences. These issues have been noted many times, and partial solutions have been found. But are we about to enter an era of deepening contrasts due to the adoption of IT? Some groups have travelled further down the road to total digital data, and this may bar access for purely technical reasons. Integrated multimedia systems—bringing together numeric, textual, pictorial and audio data—for conveying information to archaeologists (cf. Makkuni Ch. 20, this volume; Rahtz, Hall

& Allen Ch. 22, this volume) may present problems due to the existing gradient in access to hardware and communication channels. These differentials, whereby archaeological groups around the world differ, at least in the short term, in their access to IT (and hence information), are actually reinforced by the technology which is supposed to help bring about the democratization of knowledge.

Even if the technology is used in the context of collaboration there are still technical and cultural problems. Karega-Munene (Ch. 3, this volume) notes the prices of books in Kenya—but not everyone can use books, even if they can afford them. Conversely, how many readers of this book could have utilized it if it had been presented on CD-ROM? Bagnall (1989) points out that, in the long term, libraries and institutions will be able to build very extensive collections of material, based on compact discs, at modest prices compared with the building up of a library collection. Against this, many individuals are facing steep threshold costs which are impossible to meet, at least in the short term, and there are other foreseeable and hidden costs (such as those involved in backing-up or transferring data to other media should one particular technology be superseded by another) which must also be faced. Gordon (1991) quite rightly asks 'how safe is your data?' For true portability, converters between standards must be envisaged, because, even in mature fields like video and television, we are still bedevilled by different standards (eg. PAL, NTSC and SECAM).

Concluding remarks

So, where are we now? Are the worries and aspirations discussed above red-herrings? Some might say that IT in archaeology is really a cosmetic which might provide prettier pictures but does little for archaeological theory (cf. Gardin 1991). Admittedly, it is hard to deny a suspicion that IT is sometimes used to store data instead of promoting understanding of them and the furtherance of learning. Others might suggest that archaeology is being damaged by over-formalizations which make the archaeological record seem more logical than it really is. Supporters of IT can argue that at last archaeologists can progress, moving beyond the world of words to the world of images, as visionaries like Leo Biek have argued for steadfastly for so many years (Biek 1974; Biek 1976; Biek 1977; Biek 1978; Biek, Hawker, Lindsey & Lupton 1982; Biek 1985).

It cannot be denied that changes in the style of working, and interacting with others, are prompted by the information technologies geared towards analysis, visualization, and publication. It remains to be seen whether there will be a corresponding change in how data are considered. Certainly, previously undreamt of volumes of data can be handled through IT, and the implication is that many more *types* of data (especially relating to shape and spatial considerations) can be scrutinized in detail. This permits a re-examination of classificatory schemes, and can also affect our recording techniques.

But there are two great dangers facing us. The first is that we will no longer recognize incomplete data, because the quantity of data presented, and the sophistication of that presentation, may blind the uncritical to the fact that only a tiny proportion of any 'reality' may have been captured; the virtual worlds that are only just being explored may become beguilingly believable if careful self-discipline is not exercised. Illusions can be created by technological manipulation of sensory perception, rather than more sophisticated thinking. The second is that the western world will once again 'contaminate' the developing countries; not with physical disease, but with mental atrophy. The perceptions and classifications of archaeological data (eg. pottery shape) often seem to be assigned a disproportionate amount of credibility when presented by a machine—a late twentieth-century version of 'the spade never lies' syndrome. The newer presentations of space by GIS systems, for instance, may preclude original formulations of large-scale views. The latter example may, of course, be turned around; many traditional archaeologists have difficulty 'reading' GIS or data visualization output, and fresh ideas are needed on how best to exploit them. The view put forward here is that a high priority should be given to spreading full knowledge of technology as fast as possible, so that ideas for its use can come from all over the world, rather than only flowing out from the creators or originators of these systems. This is the most promising way to ensure that full discussion takes place and the technology is used to the best advantage for the study of archaeology.

Acknowledgements

Our thanks must go to Jane Hubert and Peter Ucko for their constructive comments on earlier drafts of this chapter.

References

Aldenderfer, N.S. 1987. *Quantitative Research in Archaeology: progress & prospects.* London: Sage.

Alvey, B. 1989. Hindsite. *Archaeological Computing Newsletter* **19,** 4–5.

Alvey, B. 1990. Graphic interpretations of stratigraphy using desktop computers. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 1–6. Winchester: IBM UK Scientific Centre.

Ammerman, A. 1989. The use of Landsat imagery for reconnaissance work in the Sahara. In *Late Prehistory on the Nile and the Sahara Basin,* Krzyzaniak, L. & Kobusiewicz, M. (eds), 19–25. Poznan: Poznan Archaeological Museum.

Anderson, K. 1991. The Geosoft program: creating fancy pictures. *Archaeological Computing Newsletter* **24,** 11–19.

Andresen, J. 1988. The status of the use of the computer in Danish archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. (ed.), 409–416. Oxford: British Archaeological Reports.

Anon. 1991. A new email list. *Archaeological Computing Newsletter* **27,** 14.

Arnold, C.J., J.H.Huggett, P.Reilly & S.Springham 1989. Mathrafal: a case study in the application of computer graphics. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 147–156. Oxford: British Archaeological Reports.

Arroyo-Bishop, D. 1991. The ArchéoDATA system—towards a European archaeological document . In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 61–69. Oxford: British Archaeological Reports.

Arroyo-Bishop, D. & M.Zarzosa 1989. Archaeological computing in south-western Europe. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. & J.Richards (eds), 319–325. Oxford: British Archaeological Reports.

Bagnall, R. 1989. Technology and the future of the scholarly community: three challenges. In *Sciences Historiques, Sciences du Passe et Nouvelles Technologies d'Information: bilan et évaluation,* Cacaly, S. & Losfeld, G. (eds), 403–412. Lille: CREDO.

Bennet, S. 1986. 'Dig'—an archaeological simulation. *Archaeological Computing Newsletter* **7,** 18–20.

Biek, L. 1974. Progress with LERNIE. In *Computer Applications in Archaeology 1974,* 59–63. Birmingham: University of Birmingham.

Biek, L. 1976. LERNIE—Phase III. In *Computer Applications in Archaeology 1976,* 65–71. Birmingham: University of Birmingham.

Biek, L. 1977. LERNIE—Phase IV. In *Computer Applications in Archaeology 1977,* 21–26. Birmingham: University of Birmingham.

Biek, L. 1978. LERNIE—Phase V. In *Computer Applications in Archaeology 1978,* 53–60. Birmingham: University of Birmingham.

Biek, L. 1985. LERNIE XIV: comparology and stereovideo. In *Computer Applications in Archaeology 1985,* 1–35. London: Institute of Archaeology.

Biek, L., B.M.Hawker, N.E.Lindsey & G.N.Lupton 1982. LERNIE VIII. In *Computer Applications in Archaeology 1981,* 8–19. Birmingham: University of Birmingham.

Biro, K. 1990. A microcomputer database system for the investigation of lithics. In *New Tools from Mathematical Archaeology,* Voorrips, A. & B.Ottaway (eds), 107–113. Cracow: Polish Academy of Sciences, Cracow Branch.

Boast, R. & D.Chapman 1991. SQL and hypertext generation of stratigraphic adjacency matrices. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 43–52. Oxford: British Archaeological Reports.

Boast, R. & Z.Tomlinson 1990. Computers in the city: archaeological information systems. *Antiquity* **64,** 662–666.

Booth, B. 1988. The SAM record—past present and future. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. (ed.), 379–388. Oxford: British Archaeological Reports.

Booth, B., S.Grant & J.Richards (eds) 1989. *Computer usage in British Archaeology.* Birmingham: Institute of Field Archaeologists.

Bruschini, S. & M.Politi 1990. The work of Conzorzio Neapolis. In *Communication in Archaeology: a global view of the impact of information technology. Volume two: national strategies, artificial intelligence, communication,* Reilly, P. & S.Rahtz (eds), 6–12. Southampton: University of Southampton.

Carr, R.M. 1991. The point of the pen. *Byte* **February,** 211–221.

Chenall, R. 1968. The impact of computers on archaeological theory: an appraisal and a projection. *Computers and the Humanities* **3,** 15–24.

Cleere, H. (ed.) 1989. *Archaeological Heritage Management in the Modern World.* London: Unwin Hyman.

Colley, S., S.J.P.Todd & N.Campling 1988. 3-D computer graphics for archaeological data exploration: an example from saxon Southampton. *Journal of Archaeological Science* **15,** 99–106.

Cooper, M. 1985. Computers in British field archaeology: the need for a national strategy. In *Current Issues in Archaeological Computing,* Cooper, M. & J.Richards (eds), 79–91. Oxford: British Archaeological Reports.

Cooper, M. & J.Richards (eds) 1985. *Current Issues in Archaeological Computing.* Oxford: British Archaeological Reports.

Desachy, B. & F.Djindjian 1991. Matrix processing of stratigraphic graphs: a new method. *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 29–37. Oxford: British Archaeological Reports.

Djindjian, F. 1990. A select bibliography of French data analysis applications in archaeology. *Science and Archaeology* **32,** 63–68.

Djindjian, F. 1991. Nouvelles tendances méthodologiques dans le traitement de l'information en archéologie. *Archeologia e Calcolatori* **1,** 9–13.

Dobson, M. 1990. What shall we teach them? Computers and archaeology undergraduates. *Archaeological Computing Newsletter* **24,** 17–19.

Doran, J. 1987. Anthropological archaeology, computational modeling, and expert systems. In *Quantitative research in archaeology: progress and prospects,* Aldenderfer, M. (ed.), 73–88. California: Sage Publications.

Doran, J. & F.R.Hodson 1975. *Mathematics and Computers in Archaeology.* Edinburgh: Edinburgh University Press.

Durliat, M., R.Saint-Jean, D.Kuentz & P.Lorimy 1990. *Saint-Guilhem-le-Desert. La sculpture du cloitre de l'abbaye de Gellone.* Montpellier: Association des Amis de Saint-Guilhem-Le-Desert.

Eiteljorg, H. 1988. *Computer-Assisted Drafting and Design: new technologies for old problems.* Bryn Mawr: Center for the Study of Architecture.

Gaines, S. 1984. The impact of computerized systems on American archaeology: an overview of the past decade. In *Informations Systems in Archaeology,* Martlew, R. (ed.), 63–76. Gloucester: Alan Sutton.

Gaines, S. 1987. A nostalgic farewell to ACA. *Advances in Computer Archaeology,* **4(Fall),** 1–4.

Gallay, A. 1989. Logicism: French archaeological theory. *Antiquity* **63,** 27–39.

Gardin, J.-C. (ed.) 1988. *Artificial Intelligence and Expert Systems: case studies in the knowledge domain of archaeology.* Chichester. Ellis Horwood.

Gardin, J.-C. 1990. The structure of archaeological theories. In *Mathematics and Information Science in Archaeology: a flexible frame work,* Voorrips, A. (ed.), 7–28. Bonn: Holos.

Gardin, J.-C. 1991. *Le Calcul et La Raison.* Paris: Éditions de l'École des Hautes Études en Sciences Sociales.

Gardin, J.-C., O.Guillaume, P.-Q.Herman, A.Hesnard, M.-S.Lagrange, M.Renaud, & E.Zadora-Rio 1988. *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology.* Chichester: Ellis Horwood.

Gathercole, P. & D.Lowenthal 1990. *The Politics of the Past.* London: Unwin Hyman.

Girdwood, A. 1988. Phototypesetting and desk-top publishing systems and archaeology. In *Computer and Quantitative Methods in Archaeology,* Ruggles, C. & S.Rahtz (eds), 295–299. Oxford: British Archaeological Reports.

Gordon, S. 1991. How safe is your data? In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 75–79. Oxford: British Archaeological Reports.

Grace, R. & C.Orton 1989. Hypercard as a teaching tool. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 327–338. Oxford: British Archaeological Reports.

Grant, E. (ed.) 1986. *Central Places, Archaeology and History.* University of Sheffield: Department of Prehistory and Archaeology.

Greene, K. 1988. Teaching computing to archaeology students. *Archaeological Computing Newsletter* **14,** 16–19.

Guio, A.D. 1991. Calcolatori ed archeologia: un progetto per gli anni '90. *Archeologia e Calcolatori* **2,** 25–78.

Guo, C.forthcoming. Perspectives on the use of information technology in the preservation of cultural remains and archaeological studies in China. *Archaeological Computing Newsletter.*

Haigh, J.G.B. 1989. Rectification of aerial photographs by means of desk-top systems. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 111–118. Oxford: British Archaeological Reports.

Hartman, J. 1988. Interpreting earth measurements by image enhancement techniques. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 223–247. Oxford: British Archaeological Reports.

Hassall, T. 1984. Publication: crisis and inaction. In *Informations Systems in Archaeology,* Martlew, R. (ed.), 142–152. Gloucester: Alan Sutton.

Hodder, I. (ed.) 1978. *Simulation Studies in Archaeology.* Cambridge: Cambridge University Press.

Hodder, I. 1989. Preface. In *The Meanings of Things: material culture and symbolic expression,* Hodder, I. (ed.), xxv–xxvii. London: Unwin Hyman.

Huggett, J. 1985. Expert systems in archaeology. In *Current Issues in Archaeological Computing,* Cooper, M. & J.Richards (eds), 123–142. Oxford: British Archaeological Reports.

Huggett, J. 1990. Programming AutoCAD for the archaeologist. *Archaeological Computing Newsletter* **25,** 18–24.

Jakobs, K. & K.Kleefeld 1991. Using public communication services for archaeological applications. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 3–7. Oxford: British Archaeological Reports.

Johnson, I. 1984a. Computer applications in Australia. *Advances in Computer Archaeology* **1 (Spring),** 1–7.

Johnson, I. 1984b. MINARK: a database system for archaeologists. *Advances in Computer Archaeology* **1(Spring),** 12–28.

Jones, R., A.Locker, J.Coy & M.Maltby (eds) 1981. *Computer Based Osteometry. Data capture user manual 1.* London: Ancient Monuments Laboratory.

Kampffmeyer, U. 1986. ARCOS: a documentation and database system for use in archaeology. *Advances in Computer Archaeology* **3 (Spring),** 1–40.

Kann-Rasmussen, L. & C.Larsen 1991. Photography and image databases: documentation at the National Museum of Denmark. *Archaeological Computing Newsletter* **24,** 1–7.

Knapp, A.B. 1986. Micros, Minark, and data processing at Pella in Jordan. *Advances in Computer Archaeology* **3(Fall),** 1–21.

Korsmo, T.B., F.Plog, S.Upham & L.Cordell 1985. An application of computers and computerized devices in the field. *Advances in Computer Archaeology* **2(Fall),** 1–43.

Layton, R. (ed.) 1989. *Who Needs the Past?* London: Unwin Hyman.

Lock, G. & C.Dallas 1990. Compact disc-interactive: a new technology for archaeology? *Science and Archaeology* **32,** 5–14.

Madsen, T. 1988. Prospects for the use of the computer in Danish archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. (ed.), 417–422. Oxford: British Archaeological Reports.

Majone, G. 1980. Policies as theories. *Omega* **8,** 151–162.

Martinez, V. & G.Lopez 1989. Computer archaeology in Spain. In *Communication in Archaeology: a global view of the impact of information technology. Volume two: national strategies, artificial intelligence, communication,* Reilly, P. & S.Rahtz (eds), 20–22. Southampton: University of Southampton.

Martlew, R. 1987. 'Digging deeper into history': a computer-based simulation pack. *Archaeological Computing Newsletter* **13,** 1–5.

Martlew, R. 1989. Pictures and Pedagogy: teaching archaeology using image archives on videodisc. *Science and Archaeology* **31,** 12–19.

Martlew, R. 1990. Videodiscs and the politics of knowledge. In *Humanities and the Computer: new directions,* Miall, D. (ed.), 39–48. Oxford: Oxford University Press.

Maytom, J. & K.Torevell 1991. Putting the public in the picture. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 21–22. Oxford: British Archaeological Reports.

McPherron, S. & H.Dibble 1987. Hardware and software complexity in computerizing archaeological projects. *Advances in Computer Archaeology* **4(Fall),** 25–40.

McVicar, J.B. & S.Stoddart 1986. Computerising an archaeological excavation: the human factors. In *Computer Applications in Archaeology 1986,* 225–27. Birmingham: University of Birmingham.

O'Brien, H. & C.Redman 1985. The microcomputer: an effective field tool? *Advances in Computer Archaeology* **2(Fall),** 43–67.

O'Neil, D. 1984. Archaeological uses of microcomputers with 'off-the-shelf' software. *American Antiquity* **49,** 809–814.

Oppenheim, C. 1984. Online information retrieval, data-base and networks. In *Informations Systems in Archaeology,* Martlew, R. (ed.), 90–98. Gloucester: Alan Sutton.

Orton, C. 1980. *Mathematics in Archaeology.* London: Collins.

Ozawa, K. 1989. Rule-based dating of artefacts. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. & J.Richards (eds), 375–386. Oxford: British Archaeological Reports.

Plog, F. & D.Carlson 1989. Computer applications for the All American Pipeline Project. *Antiquity* **63,** 256–267.

Polanyi, M. 1958. *Personal Knowledge.* London: Routledge & Kegan Paul.

Powell, D. 1984. Publishing in the electronic age—an overview of development in the publishing industry. In *Informations Systems in Archaeology,* Martlew, R. (ed.), 126–130. Gloucester: Alan Sutton.

Powlesland, D. 1983. Pots, pits and portables. *Practical Computing* **June,** 144–146.

Powlesland, D. 1985. *Archaeology and Computers.* London: Council for British Archaeology.

Presland, G. 1984. A cautionary tale: computerization of the Victoria archaeological survey sites register. *Advances in Computer Archaeology* **1(Spring),** 7–11.

Quattrin, C. 1990. Della conoscenza alla tutela degli affreschi di Pompei: un sistema esperto al servizio della comunita archeologica. In *Communication in Archaeology: a global view of the impact of information technology. Volume two: national strategies, artificial intelligence, communication,* Reilly, P. & S.Rahtz (eds), 47–53. Southampton: University of Southampton.

Rahtz, S. 1986. AIE—archaeological information exchange. *Archaeological Computing Newsletter* **8,** 10.

Rahtz, S. 1989. Reflections on an archaeological information exchange. *Archaeological Computing Newsletter* **16,** 10–15.

Rains, M. 1989. Computerising the plan database: the Aegis system. *Archaeological Computing Newsletter* **21,** 1–2.

Rakatoarisoa, J.-A. 1990. Unpublished paper presented at the Second World Archaeological Congress, Barquisimeto. Barquisimeto: World Archaeological Congress 2.

Ravetz, J. 1971. *Scientific Knowledge and its Social Problems.* Oxford: Clarendon Press.

Reilly, P. 1985. Computers in field archaeology: agents of change? In *Current Issues in Archaeological Computing,* Cooper, M. & J.Richards (eds), 63–78. Oxford: British Archaeological Reports.

Reilly, P. 1989. Data visualization in archaeology. *IBM Systems Journal* **28(4),** 569–579.

Reilly, P. & J.Richards 1988. New perspectives on Sutton Hoo: the potential of 3-D graphics. In *Computer Applications in Archaeology 1987,* 173–185. Oxford: British Archaeological Reports.

Reilly, P., J.D.Richards & A.Walter 1988. Graphics at Sutton Hoo: the potential of 3-D graphics. *Bulletin of the Sutton Hoo Research Committee* **5,** 24–27.

Richards, J. 1985. Standardising the record. In *Current Issues in Archaeological Computing,* Cooper, M. & J.Richards (eds), 93–102. Oxford: British Archaeological Reports.

Richards, J. 1988. ARCH-HELP: computer-aided help for archaeology students. *Archaeological Computing Newsletter* **15,** 9–13.

Richards, J.D. 1986. *Computer usage in British Archaeology.* Birmingham: Institute of Field Archaeologists.

Richards, J.D. 1991. Terrain modelling, deposit survival and urban archaeology. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 175–82. Oxford: British Archaeological Reports.

Roberts, M.L. 1990. The Franklin Institute: building the 'smart' museum of science and technology. *Unisphere* **10(4),** 34–36.

Roosevelt, A. 1987. Microcomputers in the lower Amazon project. *Advances in Computer Archaeology* **4(Fall),** 1–53.

Ruggio, A.Z. 1991. Sperimentazione dell'informatica nella didattica dell'arceologia. *Archeologia e Calcolatori* **2,** 19–23.

Ryan, N. 1988. Browsing through the archaeological record. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 327–334. Oxford: British Archaeological Reports.

Ryder, M. 1986. The use of a digitizing tablet to measure fibre diameters in textile remains. *Archaeological Computing Newsletter* **7,** 21–24.

Sabloff, J. (ed.) 1981. *Simulations in Archaeology.* Albuquerque: University of New Mexico.

Scollar, I. 1989. The Bonn seriation and clustering package 3.1. *Archaeological Computing Newsletter* **16,** 23–25.

Scollar, I., A.Tabbagh, A.Hesse & I.Herzog 1990. *Archaeological Geophysics and Remote Sensing.* Cambridge: Cambridge University Press.

Shennan, S. 1988. *Quantitifying Archaeology.* Edinburgh: Edinburgh University Press.

Smith, F. & D.Harris 1990. The UK Science and Engineering Research Council and scientific archaeology. In *Communication in Archaeology: a global view of the impact of information technology. Volume three: late papers,* Reilly, P. & S.Rahtz (eds), 55–57. Winchester: IBM UK Scientific Centre.

Stančič, Z. 1989. Computervision: producing intra-site plans. *Archaeological Computing Newsletter* **20,** 1–10.

Stephen, D. 1987. Personal computers in the field—a tool for modern times? *Advances in Computer Archaeology* **4(Fall),** 5–24.

Stewart, W.K. 1991. Multisensor visualization for underwater archaeology. *IEEE Computer Graphics and Applications* **March,** 13–18.

Stone, P. & R.MacKenzie (eds) 1990. *The Excluded Past: archaeology in education.* London: Unwin Hyman.

Templeton, L. 1990. Archaeological illustration and computer aided design. *Archaeological Computing Newsletter* **24,** 1–8.

Trigger, B.G. 1989. *A History of Archaeological Thought.* Cambridge: Cambridge University Press.

Upham, S. (ed.) 1979. *Computer Graphics in Archaeology: statistical cartographics applications in archaeological contexts.* Arizona: Arizona State University.

Voorrips, A. 1987. Formal and statistical models in archaeology. In *Quantitative Research in Archaeology: progress and prospects,* Aldenderfer, M. (ed.), 61–72. London: Sage.

Voorrips, A. & B.Ottaway (eds) 1990. *New Tools from Mathematical Archaeology.* Cracow: Polish Academy of Sciences, Cracow Branch.

Wandibba, S. 1990. Archaeology and education in Kenya. In *The Excluded Past: archaeology in education,* Stone, P. & R.MacKenzie (eds), 43–49. London: Unwin Hyman.

Weiss, A. 1989. The COMPASS system. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. & J.Richards, (eds), 295–318. Oxford: British Archaeological Reports.

Whallon, R. 1972. A new approach to pottery typology. *American Antiquity* **37,** 13–33.

Wilcock, J. 1989a. Computer archaeology in China. *Archaeological Computing Newsletter* **17,** 2–7.

Wilcock, J.D. 1989b. Review Paper: over thirty years' application of the computer in archaeology. *Science and Archaeology* **31,** 35–44.

Zarzosa, M.T.L. & D.Arroyo-Bishop 1989. New directions for archaeological computing in south-western Europe (France, Spain and Portugal). In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 319–326. Oxford: British Archaeological Reports.

2

The impact of information technology on the archaeology of southern and eastern Africa— the first decades

P.SINCLAIR, M.KOKONYA, M.MENESES & J.-A.RAKATOARISOA

Introduction

For many of us the 1980s have seen sweeping changes in the way in which we approach our archaeological data. This has in no small measure been due to the impact of information technology (IT) upon the human and natural sciences and archaeology as a data intensive discipline has changed accordingly.

Early IT applications in eastern and southern Africa involved both the development of museum-based site recording systems and also academic interest in statistical analysis and comparison of archaeological assemblages. By the 1970s certain countries of the region, such as Zimbabwe and Kenya, possessed well curated archaeological site indexes. That of Zimbabwe was modelled as early as the 1950s on the Swedish national site register (Summers *pers. comm.*). The further development of the comparative aspects of the site records owed a great deal to the still unsurpassed *Atlas of African Prehistory* (Clark 1967). To the best of our knowledge the first computer-based archaeological site index in eastern or southern Africa was that of the Zambian Antiquities service which was operational in the mid 1970s. In South Africa a centre at the South African museum in Cape Town focused upon rock art recording and slowly the examples set further to the north became incorporated into the most powerful archaeological establishment in the continent. In eastern Africa museum-based site recording systems have benefited considerably from a degree of standardization through the adoption of the SASES site recording system, which is the current regional standard.

The 1980s saw the arrival of microcomputers (notwithstanding their cost) throughout the region. An example of database possibilities using simple microcomputer systems is the c. 5000 sites registered in the Zimbabwe site index (Sinclair 1987). This was built using an MSDOS Knowledgeman database system and includes late stone age and farming community sites as well as all recorded stone enclosures. A similar system, using dBase IV, was developed for Mozambique by Ricardo Duarte *(pers. comm.)*. Mention should also be made of efforts after 1987 in the regional research programme on Urban Origins in East Africa which incorporates nine east African countries (Somalia, Kenya, Comores, Madagascar, Mozambique, Tanzania mainland and Zanzibar, Zimbabwe and Botswana). This facilitated the spread of microcomputer technology to some archaeologists, most of whom are working in areas of the continent and the offshore islands which had little or no previous access to IT. A project specialist meeting in Mombasa further widened discussion on such topics as standardization and exchange of information and a set of flexible guidelines was agreed upon (Sinclair 1988). This involved not only site recording systems but also radiocarbon date reporting. The current situation is that IBM compatible microcomputer systems with at least a 286 level processor are established in each of the participating countries, and Macintosh systems are also present in Kenya and Madagascar. Widely distributed software systems include Microsoft Works, dBase III+ and IV, and Minark database systems. CAD systems (Autocad) are just beginning to be established in east African archaeology; the Idrisi geographic information system was used successfully at the postgraduate researcher training course in Uppsala in 1991.

Quantitative research in archaeology

Academic interest in IT applications in eastern and southern Africa seem to flower rapidly but are often of shorter duration than the curatorial interests which respond to an ever present and increasing need. Financial constraints pose severe limits on the rate of development of systems. Research applications can be viewed succinctly as falling within the four major archaeological paradigms outlined by Clarke (1968). Much of the early South African work was firmly aimed at morphological trait analysis and the comparison of stone tool assemblages. At the University of the Witwatersrand in Johannesburg a number of mainframe based statistical applications, particularly the diachronic analysis of stone tool assemblages as at the Cave of Hearths, had been carried out under the direction of Mason (Mason 1971). In the Cape, H. and J. Deacon developed a new approach to late stone age stone tool typology and this was at least matched in the very considerable typological developments in Kenya and Tanzania by Kleindienst and Isaac among others. It would be interesting to know more of the role of IT in these developments. Certainly by the time of Manheim, Parkington and Robey’s correspondence analysis of stone tool assemblages of the western Cape, published in 1984, computer aided applications were well established at the university of Cape Town, where a unit for spatial archaeology was opened. In east Africa a series of scholars from the USA, notably Nelson and coworkers, were responsible for introducing microcomputer based field recording techniques.

The 1970s and early 1980s also saw the extension of the morphological trait analysis into areas which had been previously little researched. In Mozambique, computer-aided trait analysis of each individual sherd in ceramic assemblages (using a mainframe-based SAS system) was applied with the explicit aim of building up the primary chronostratigraphic framework in four areas of the country. This saw the early application of correspondence analysis (1981/2) at the Zimbabwe tradition stone wall enclosure of Manyikeni (Sinclair 1987). In a stimulating, though somewhat controversial, multi-variate analysis, use was made of computer-aided trait comparisons of east African neolithic ceramic assemblages by Collett & Robertshaw (1983). In Scandinavia considerable effort was put into the ceramic typology and analysis of the ceramics *within a context of other finds* from the joint Scandinavian expedition to Nubia (Sinclair & Troy 1989).

The role of information technology

What then has been the role of IT in these developments in eastern and southern Africa? In retrospect when considering one of the great debates between the different approaches to archaeological typology as propounded in Spaulding (1953) and Doran & Hodson (1975), the role of IT seems to have favoured the former approach. Repeated analysis of different aspects of artefact assemblages, and the creation of groupings relevant to the problem at hand, are certainly strongly favoured in IT applications, but are not limited to them, as the typological work of Soper and Huffman illustrates. In eastern and southern Africa the dominance, and indeed the relevance, of the morphological paradigm was challenged from a number of other directions.

From an ecological perspective, building upon previous work by Marker and Jones in the eastern Transvaal, Hall (1981) broke new ground with an application of multi-dimensional scaling to locational parameters of iron age sites on the Natal coast and integrated the results into a model of changing settlement systems. Increasingly, the concept of statistical sampling was being discussed, even if its application was often found wanting. In addition, obtaining the knowledge required of the range of biotic and abiotic factors necessary for successful ecological modelling was a daunting task for archaeologists trained in a humanist tradition. Nevertheless, osteological studies developed particularly in South Africa and Kenya, but also spread to Malawi, Zimbabwe and Zambia. An early IT application using correspondence analysis for comparing 50 osteological collections was carried out in the Transvaal (Mason 1983).

In the Orange Free State extensive work based upon a geographical paradigm was carried out on late stone age Smithfield tradition sites by Sampson and coworkers (Sampson 1984). The site corpus was not based on statistical sampling but on exhaustive field-walking in relatively open landscape. Detailed computer-assisted distribution and density maps were produced and the first steps taken at relating site distributions to environmental parameters. Further north in Zimbabwe Sinclair, working from the database described above, subjected sites of different phases to spatial analysis. Density mapping using Lundmark’s DCurve method which permits the selection of relevant scales for mapping spatial domains was applied to sites of different phases identified on the basis of ceramics and architectural traits. In a move towards developing gravity modelling this work included the first archaeological application of fuzzy set cluster analysis (Sinclair 1987).

Considerable potential exists in arid areas such as Botswana with its well developed settlement hierarchy (Denbow 1984) for application of these techniques. Some of the potential for IT applications in archaeological survey has been shown in an extensive survey of the Juba valley in Somalia. The utility of further IT applications, particularly using digital data from SPOT satellite imagery, is currently being investigated in both areas as part of the Urban Origins in East Africa project.

At the intra-site level, the 1970s and early 1980s saw intensive investigation of stone age sites such as Border Cave, Die Kelders and Robberg in South Africa, Kalambo Falls in Zambia, and the famous early stone age sites of the Kenya Tanzania rift *without* the application of IT. With regard to the later periods of the archaeological record, extensive excavation of iron age sites was undertaken (eg. Broederstroom in South Africa (Mason 1981), Leopards Kopje in Zimbabwe (Huffman 1974), Shanga in Kenya (Horton 1984)) with none or only minimal use of IT. Few excavations have used statistical random sampling techniques, with the exception of that by Bisson & Horne (1974) in Zambia, and at Manyikeni in Mozambique. These latter approaches broke away from the aim of total excavation, and started a trend in down-sizing the scale of excavation while maintaining, and indeed increasing, the spatial coverage of the investigation. Early work on data from Kenya by Ammerman, Gifford & Voorrips (1977) and the SURFACE II applications by Zubrow & Harbaugh (1977) have clearly shown the utility of this approach, as did the widely read experience of Flannery (1976) and co-workers in Meso-America.

In Mozambique as early as 1977, sampling units of one-metre squares chosen randomly from stratified blocks were excavated and the contents of each recorded in a database, initially with SAS on an IBM mainframe but more recently using SPSS, Knowledgeman and dBase III+ and IV under MSDOS (Sinclair 1987). In Tanzania, contour mapping techniques have been used to plot the intra-site distribution of iron working remains. Apart from basic distribution maps a number of interpolation techniques have been applied. At the University Campus site in Maputo in 1984 the spatial distribution of different categories of finds was interpolated *prior to extensive excavation* using SURFACE II, and for the first time a predictive model based upon an interpolated sample was tested in the field. This has led to the establishment of an informal, though nevertheless useful, concept of 'turn-around-time'—the length of time taken to set up empirically based predictive hypotheses, and to refute them empirically in the field. We are approaching a degree of interactivity with a wide range of archaeological data which ten years ago would have been thought impossible. While this may be true, are these 'toys' just for the rich? Is this not just another expression of methodology masquerading as theory, or imperialist manipulation of science? How many east African archaeologists have access to this new means of production? Karega-Munene (1992) gives one perspective on this situation.

Notwithstanding these problems, surprising opportunities can occur. One of the first attempts in Mozambique, for instance, at developing computer-aided archaeological applications was stimulated by the availability of a Cuban version of a Bulgarian computer with notable similarities to an IBM mini-computer, a gift from Fidel Castro to President Samora Machel of Mozambique in the late 1970s. Another example comes from Madagascar, where information technology is not really a new concept, as different government departments have been using it for several years, and colleagues from scientific departments (eg. mathematics, physics, etc.) have been familiar with it since the 1960s. There was also experience in the social sciences, but this mainly involved building databases and performing statistical analysis, with access limited to a few 'illuminated ones'.

The big change came with the use of personal computers. People wanted to have their own computer and it is possible to talk about a true fashion for this new product. However, with the exception of the people who already had a practical knowledge of computers, very few Madagascan archaeologists have made the effort to acquire any training in the use of computers. Most people have been happy to raise many different questions and compare performances of computer makes, without really having paid much attention to their real needs.

The department at Antananarivo in Madagascar only started to work with computers in 1985 thanks to two Osborne CP/M systems given by colleagues in Duke University, USA. The problem of deciding which hardware to choose did not have to be faced from scratch. This equipment did not have the performance of the Toshiba MSDOS compatibles now favoured by archaeologists in east Africa, but they performed an invaluable service by helping to start training staff in information technology. It must be stressed that equipment must be chosen according to the reality of the country, with respect to maintenance and the possibility of finding new parts. The decision was taken in Madagascar to use portables that can be taken, in case of a breakdown, to a country where it is easy to repair them. Preference is given to mains-powered equipment but with a battery to cope with the instability of the electric current which can at any moment cause breaks or peaks in voltage. In this way, one avoids buying a stabilizer and a modulator which are expensive, and are not enough to protect the computer from high variations in voltage.

In the past, it was necessary to acquire both MSDOS and Apple Macintosh systems in order to have flexibility for exchange of files with colleagues from other countries. This precaution is no longer justified since Apple have made their Macintosh able to read MSDOS disks.

Training

Before any work of the kind described above can be undertaken, the human factor has to be taken into account. Personnel have to be psychologically and technologically prepared to use new equipment; in Madagascar, personnel are trained directly, familiarizing them with each software package in use, and correcting the errors as they appear. A theoretical general introduction on the basic concepts of information technology is made available to researchers.

Although within each institution it is becoming more and more difficult to have any doubts about the use of computers, this is not however yet the unanimous opinion of all people working in the field in east Africa. There remain a number of obstacles which hinder the development of a real information technology programme for data in the social sciences.

Since the early 1980s with the consolidation of a genuinely regional frame of reference spanning the whole of eastern Africa, a new context for the discussion of the role of IT has been made available—a human network to facilitate its spread *and democratization*. One example is the current programme on Urban Origins in East Africa mentioned above (Sinclair 1988; Sinclair & Wandibba 1988; Sinclair & Rakotoarisoa 1990). Of the nine countries involved in this initiative only archaeologists from Kenya and Mozambique had previous experience of IT. As a part of its work, the project has established IBM-compatible MSDOS systems in each of the participating countries. In addition, the first east African to obtain postgraduate training in IT has recently graduated (a summary of his work is presented in this chapter and in Kokonya 1990). In each country technicians need to be introduced to the new technology and trained in word-processing and the setting up of databases. These needs are much wider than can be catered for in the present project. Courses to develop basic statistical and analytical skills have, however, been initiated and will be intensified during the remainder of the project.

On the research level twenty-two PhD candidates are studying different aspects of urbanism in eastern Africa. Their research is subsumed under four themes, which focus on chronology, inter-site spatial relationships, intra-site organization, and ethnohistorical interpretation. A beginning at building up a locational database on urban and associated sites in eastern Africa has been made using Minark and Knowledgeman database systems. Oracle applications are being considered but as yet are not justified by present levels of available data. More than 600 radiocarbon dates have been recorded from the region.

For the inter-site analyses, six areas of east Africa have been focussed upon (Kenya coast and interior, Comores, Madagascar highlands, Zimbabwe, Botswana). 17000 hillforts from highland Madagascar are being treated in a similar way to the data from the Zimbabwe Plateau (Sinclair 1987). Here again, particular needs are the identification of hierarchically nested spatial domains in the data, and relating these to environmental units. The aim of this work is to provide the spatial context within which to view the development of urban zones and town centres in the region as a whole. Systems based on 80386 processors, incorporating an IBM PS/2 80 and digitizer, have been set up in Uppsala, Sweden, to deal with this data. The kernel of the software programs is the Rockware set of programs. SURFER is also in use particularly for field applications. Currently the first attempt at a Graphical Information System (GIS) application using data from the Comores is under way. Detailed site location data from the surveys of Wright and Verin among others on the Comores have been compiled and are currently being analysed with the Idrisi GIS system.

Intra-site analysis

At the intra-site level an ambitious programme of nineteen separate excavations is planned or currently being implemented by the participants in the Urban Origins project. The main thrust of this work is away from the town centres characterized supposedly by stone architecture and towards incorporating the more extensive mud built quarters into the analytical frame of reference. In order to achieve this aim and to ensure that individual researchers are not drowned in new data, sampling strategies which favour a widespread but downsized sampling unit have been extensively applied. Many of the sites we are dealing with are from 3–5 hectares in area but some foci of interest, such as Mahilaka in northern Madagascar and the surroundings of Great Zimbabwe, measure up to 100 and 700 hectares respectively. Definition of the extent of sites and their internal stratigraphy is a key element in the understanding of urban growth along the coasts of eastern Africa as well as in the interior.

To facilitate these approaches a micro-coring technique was developed as part of the project. A portable motorized boring machine with a 50mm screw auger was tested at Mtwapa, Kenya against a known 3.5m stratigraphy. Each of the archaeological units recognized in the trench were clearly visible in the cores. Twenty-two cores were drilled in the vicinity and basic stratigraphic interpolation was carried out in the field using SURFER. In addition, soil samples were collected and tested for phosphates with positive results by the archaeological research laboratory of Stockholm University.

Since then the technique has been refined considerably and the site of Mbashile on the island of Ngazidja (part of the Comores Archipelago) will be used to illustrate briefly some aspects of the technique. The site was previously recorded by Wright and a profile along part of the sea shore was dated from the 9th to the 15th centuries with a hiatus from the 11th to the 12th centuries. Fifty-five cores were drilled to a maximum depth of three metres although normally they were terminated by two metres. Each separate unit was recorded by texture, colour and inclusions and the information incorporated into a Knowledgeman database. A Harris matrix of colour equivalents was built, and this greatly facilitated the overall stratigraphic interpretation. This framework was very useful for incorporating cultural finds as well as the results of the SPOT test for phosphate, the values of which are produced in the field. These indicate the possibility of two occupation levels separated by a sterile soil horizon, the first on ceramic evidence dating to the end of the first millennium AD and the second from c. 13th–15th centuries. The first abandonment of the site could well coincide with increased erosion resulting from a depletion of vegetation on the slopes of the nearby volcanic outlier. This interpretation is based on just over of a square metre of deposit and is of course open to criticism from a number of points of view. It should, however, be seen as providing a predictive hypothesis for future investigation of the site.

More sophisticated use of the Rockware modules has been made by project participants helping with the excavation of the Swedish medieval town site of Birka. Here superimposed stratigraphic units were interpolated using the Gridzo module, sections drawn and volumetric estimates provided for each of the units (Sinclair 1990).

Similar drilling approaches are currently under way in Madagascar, Kenya, Zanzibar and in Zimbabwe. Resistivity techniques have been successfully applied in Madagascar and magnetometry in Tanzania as well. Particularly clear results have been obtained by Rakotovololona in Madagascar at the highland site of Ankadivory and by Radimilahy in Mahilaka on the northern coast. In Zimbabwe, at Great Zimbabwe, a research programme involving ground survey and sampling some 700 hectares and the coring of an approximately two-kilometre strip through the site has been undertaken. In addition to spot and colour tests, magnetic susceptibility measurements have been introduced and these provide another insight into the area around the mainsite. Additional use has been made of the Scollar Harris matrix program (Herzog & Scollar 1991) to correlate results from more than 200 cores. These results have been incorporated into a solid model of six stratigraphic units covering an area of twenty hectares in the centre of the site using Gridzo modules from the Rockware system.

Observations on the introduction of information technology

The introduction of IT into archaeological programmes in east Africa has not been entirely unproblematic. Indeed, a number of obstacles have been encountered. The majority have their origins in such critical areas as received ideology, price considerations, technophobia, and a basic ignorance of the fundamentals of IT in many cases.

Negative reactions are by no means unknown. Certain people still think that the development of IT is a phenomenon which risks getting out of control. Madagascar, for instance, experienced a period when its government wanted to instigate strict controls on the importation of computer equipment in order to try and standardize systems. An official certificate was required before any computer could be imported. This procedure was abandoned, and the market is now completely open. In addition, the use of computers was considered as being a reserved domain for mathematicians; this is no longer true but the idea is still prevalent.

Economics are a major consideration in the introduction of IT. Many people consider the use of computers as a luxury in the case of developing countries, although their prices are often only slightly higher than those of typewriters.

Moreover, many people remain apprehensive of the new technologies. A psychological block to the use of computers still remains and is probably the most difficult obstacle to overcome. A fear of change is the most likely cause of this apprehension. Computers change our habits; a researcher who used to spend most of his or her time going through files instead of thinking about the subject is now forced to do the opposite. The conversion effort required seems to be out of reach for many researchers. This generates a real fear of computers.

In common with many other African countries, Madagascar suffers from a real lack of IT literature, with few magazines and periodicals being available. Most people try to obtain such material through networks of friends. This lack of information can result in choices being made which do not suit the requirements (eg. the purchase of a machine with only 256k of RAM, which excludes the use of most packages).

The solution does not consist simply in getting more money or equipment. Collaborations are needed at all levels, but firstly with the introduction of basic training in software exploitation. Developing systems comes later in the process of technology transfer. For only when enough people understand the fundamentals of the technology being used can we start to consider seriously, and attempt to solve, the specific IT needs of individual countries. The great strength of IT is the flexibility which makes it possible to adjust them to every situation and user.

There is no doubt that computer graphics and image processing and the other techniques that have been discussed in this book (Fletcher & Spicer 1992; Molyneaux 1992; Reilly 1992; Wood & Chapman, (with contributions by Delooze & Trueman) 1992) are going to play an increasingly important role in archaeological research. Clearly, these techniques offer many new possibilities in the way data is perceived. Certainly, new ideas can be generated and tested easily and in ways which hitherto were unavailable. These new methods of looking at our data, however, raise further problems and challenges which are yet to be resolved. Several things stand in the way of the wider usage of advanced techniques in African archaeology. Of course, not all of these things are unique to the African situation. For instance, computer graphics and image processing require, in the majority of cases, special purpose hardware which is very expensive, putting them out of reach of most archaeologists. Graphics programming, though a field which is very exciting, is demanding because of some the elusive concepts which are involved. ‘Direct-manipulation user interfaces’ have only provided a partial answer to this problem. There is, as yet, no equivalent, for example, of the higher level abstractions like ‘Fourth Generation’ tools that are to be found in the database world. Standardization in computer graphics has been another problem which has prevented the development of portable graphics software. This problem has partly been tackled through recent work on standards like the GKS and GKS-3D specifications, and other de-facto standards like PostScript.

In fact, although some advanced techniques show great potential, they are not fully tested or accepted in the world of archaeology generally. The complexity of many of these methods is a common obstacle to their wider utilization. For the more sophisticated the method, and the higher the skills required to harness it, the lower the rate at which the techniques are likely to become incorporated into working practices. If used blindly, these advanced data exploration methods can be very deceiving. Spicer (1988) discusses some of these problems and issues associated with the perception of archaeological information using graphical techniques.

The new methods have surely demonstrated some inadequacies and failings in the recording of primary data and have thus raised several methodological issues. These offer fundamental challenges to our recording strategies and may require some radical revisions in the way we operate. For example, we often record data with no particular analysis in mind, but in the hope of some future researcher finding use for it. One can see the interpretive advantages of three-dimensional visualizations of archaeological data; yet most of our recording systems and strategies are still based on a two-dimensional space and thus are very poor representations of our data. A further point should be made concerning the way we obtain our primary data. The methods developed so far for data acquisition have not made this task any less tedious and time consuming. The archaeologist's time is not optimized for research and analysis. Rather data acquisition and ordering are the principal activities. In a number of situations, and particularly with survey data, data acquisition (eg. coring) would clearly benefit from some automation.

One other major obstacle to the wider utilization of the new tools and methods, particularly in the less technologically developed parts of the world, is communication. Lack of access to international networks and information systems has often held back the development and use of computer-based tools, because it takes far longer for new techniques to percolate through to those without access to the networks.

Some of these problems and challenges may be overcome with time. The high cost of graphics facilities, for example, will eventually come down with the rapid developments currently in evidence. We should avoid the situation where computer technology creates barriers between archaeologists in the so called developed and developing worlds. A final word of caution: one should not adopt new technologies or methods simply because it is fashionable to do so; one must critically assess their role and suitability to the archaeological problems. There is a risk of over-emphasising the possibilities of these methods, getting lost among the technicalities, and forgetting the initial archaeological problems. New tools should enhance human creativity rather than stifle the development of new ideas.

Conclusions

Without doubt, over the last ten years IT applications have multiplied in eastern and southern Africa. The potential of this technology has barely been addressed. Many problems remain, not least the costs of equipment, the small numbers of systems available, the isolation of scholars and even technophobia of researchers trained in a humanist tradition. All of these factors and many more act against the hoped for democratization of IT. Nevertheless, given the progress in the region over the last five years in consolidating a basis of regional cooperation, it should be possible to enlarge the field of IT in eastern Africa to permit a wider range of scholars to contribute their talents to solving problems which face us all.

References

Ammerman, A.J., D.P.Gifford & A.Voorrips 1977. Towards an evaluation of sampling strategies: simulated excavation of a Kenya pastoralist site. In *Simulation Studies in Archaeology* Hodder, I. (ed.), 123–132. Cambridge: Cambridge University Press.

Bisson, S. & S.Horne 1974. Probability sampling techniques for investigation of African iron age sites. *Azania* **9,** 87–104.

Clark, J.D. (ed.) 1967. *The Atlas of African Prehistory.* Cambridge: Cambridge University Press.

Clarke, D.L. 1968. *Analytical Archaeology.* London: Methuen.

Collett, D. & P.Robertshaw 1983. Pottery traditions of early pastoral communities in Kenya. *Azania* **18,** 107–123.

Denbow, J.R. 1984. Cows and kings: a spatial and economic analysis of a hierarchical early iron age settlement system in eastern Botswana. In *Frontiers: Southern African Archaeology today,* Hall, M., Avery, G., Avery, D.M., Wilson, M.L., & Humphreys, A.J.B. (eds). Oxford: British Archaeological Reports.

Doran, J. & F.R.Hodson 1975. *Mathematics and Computers in Archaeology.* Edinburgh: Edinburgh University Press.

Flannery, K. 1976. *The Meso-American Village.* New York: Academic Press.

Fletcher, M. & D.Spicer 1988. Experiments with gridded survey data. In *Computer and quantitative methods in archaeology 1987,* Ruggles, C.L.N. & S.Rahtz (eds), 201–220. Oxford: British Archaeological Reports.

Fletcher, M. & D.Spicer 1992. The display and analysis of ridge-and-furrow from topographically surveyed data. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 10. London: Routledge.

Hall, M. 1981. *Settlement Patterns in the Iron Age of Zululand.* Oxford: British Archaeological Reports.

Herzog, I. & I.Scollar 1991. A new graph theoretic oriented program for Harris Matrix analysis. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & Rahtz, S. (eds), 53–60. Oxford: British Archaeological Reports.

Horton, M. 1984. *The Early Settlement of the Northern Swahili Coast.* PhD thesis, University of Cambridge, Cambridge: University of Cambridge.

Huffman, T.N. 1974. *The Leopard's Kopje Tradition.* Salisbury, Zimbabwe: National Museums and Monuments of Rhodesia.

Karega-Munene 1992. Dissemination of archaeological information: the East African experience. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 3. London: Routledge.

Kokonya, M. 1990. Data exploration in archaeology: new possibilities and new challenges. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 2: national strategies/artificial intelligence/communication.* Southampton: Department of Electronics and Computer Science, University of Southampton.

Manheim, A., J.E.Parkington & T.S.Robey 1984. Stone tools and sandveld settlement. In *Frontiers: southern African archaeology today,* Hall, M., G.Avery, D.M.Avery, M.L.Wilson & A.J.B.Humphreys (eds). Oxford: British Archaeological Reports.

Mason, R.J. 1971. *Multivariate Analysis of Cave of Hearths Middle Stone Age Artifact Assemblages.* Johannesburg: Department of Archaeology, University of the Witwatersrand.

Mason, R.J. 1981. Early iron age settlement at Broederstroom 24/73. Transvaal, South Africa. *South African Journal of Science* **77(9),** 401–410.

Mason, R.J. 1983. Cluster and correspondence analysis of iron age faunal assemblages from the southern western and eastern Transvaal. *South African Journal of Science* **79,** 189–203.

Molyneaux, B.L. 1992. From virtuality to actuality: the archaeological site simulation environment. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 19. London: Routledge.

Reilly, P. 1992. Three-dimensional modelling and primary archaeological data. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 12. London: Routledge.

Reilly, P. & S.Rahtz (eds) 1992. *Archaeology and the Information Age.* London: Routledge.

Ruggles, C.L.N. & S.Rahtz (eds) 1988. *Computer and Quantitative Methods in Archaeology 1987.* Oxford: British Archaeological Reports.

Sampson, G. 1984. Site clusters in the Smithfield settlement pattern. *South African Archaeological Bulletin* **39,** 5–23.

Sinclair, P. 1987. *Space, Time and Social Formation.* PhD thesis, University of Uppsala, Uppsala: University of Uppsala.

Sinclair, P. 1988. *Urban Origins in Eastern Africa: the Mombasa specialist workshop.* Urban Origins in Eastern Africa Project Working Papers 3. Stockholm: Central Board of National Antiquities.

Sinclair, P. 1990. *Micro Stratigraphic Investigations at the Medieval Town Site of Birka.* Urban Origins in Eastern Africa Project Working Papers 5. Stockholm: Central Board of National Antiquities.

Sinclair, P. & J.A.Rakotoarisoa 1990. *Urban Origins in Eastern Africa: proceedings of the Madagascar Workshop.* Urban Origins in Eastern Africa Project Working Papers 4. Stockholm: Central Board of National Antiquities.

Sinclair, P. & L.Troy 1989. A multivariate analysis. In *Middle Nubian Sites,* Säve-Söderbergh, T. (ed.), 273–290. Partille, Sweden: Paul Osröm Editions.

Sinclair, P. & S.Wandibba 1988. *Urban Origins in Eastern Africa: proceedings of the 1988 Nairobi Workshop.* Urban Origins in Eastern Africa Project Working Papers 1. Stockholm: Central Board of National Antiquities.

Spaulding, A.C. 1953. Statistical techniques for the discovery of artefact types. *American Antiquity* **18,** 305–313.

Spicer, D. 1988. Computer graphics and the perception of archaeological information. In *Computer and Quantitative Methods in Archaeology 1987,* Ruggles, C.L.N. & S.Rahtz (eds), 187–200. Oxford: British Archaeological Reports.

Wood, J. & G.Chapman (with contributions by K.Delooze & M.Trueman) 1992. Three-dimensional computer visualization of historic buildings—with particular reference to reconstruction modelling. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 11. London: Routledge.

Zubrow, E.W. & J.W.Harbaugh 1977. Archaeological prospecting: kriging and simulation. In *Simulation Studies in Archaeology,* Hodder, I. (ed.), 109–122. Cambridge: Cambridge University Press .

3

Dissemination of archaeological information: the east African experience

KAREGA-MUNENE

Introduction

Knowledge of east African archaeology began in the 1920s. This was as a result of the work of the late Louis Leakey and, later, Mary Leakey. Their contribution during the first half of this century and through the 1960s laid a fairly firm foundation for the practice of archaeology in east Africa. Consequently, east Africa is today widely known for its rich archaeological information since it has the longest record of occupation by humans and their ancestors. The evidence for this—which is beyond the scope of this chapter—has been well documented (*inter alia*, Isaac & McCown 1976; Johanson&Edey 1981; Gowlett 1984; Ciochon & Fleagle 1985; Tobias 1985). Therefore, it will suffice to make just a few observations here.

To begin with, the east African archaeological record dates as far back as the beginning of the Miocene epoch, between 23 and 18 million years ago. The evidence for this is abundant in the region. This is because the earth movements associated with the Great Rift Valley have repeatedly created sedimentary basins in which hominid fossils and ancient archaeological sites have been buried and preserved with a minimum of disturbance. The subsequent break up and elevation of parts of this formation by earth movements have exposed them to erosion, leading to exposure of artefacts and fossils. Recent archaeological sites are also exposed by erosion because of the nature of the topography. In addition, continual volcanic activity in the region provides a basis for radiometric dating of both palaeontological and archaeological finds especially using the potassium-argon technique (K-Ar). Elsewhere in Africa K-Ar is not applicable; therefore, ancient finds made there are dated by correlating them with similar finds from east Africa.

This chapter aims to examine how this information has been disseminated amongst the region’s archaeological community as well as amongst the general public. Thus, it evaluates the merits and demerits of conventional methods of making archaeological data available. Additionally, it evaluates the import of the argument that the growing availability of computer networks and other technologies is likely to mean that communication will be increasingly available to archaeologists in all parts of the world, with specific reference to the east African situation.

Research and publications

For a long time now east Africa has been witnessing much archaeological research each year. This has, and continues to, involve, both resident and foreign archaeologists and their students. Virtually all the research findings are published overseas in specialised journals, monographs and textbooks. The only exceptions are the dissertations and theses written by local students and the journal *Azania,* which is published by the British Institute in eastern Africa, and more recently the newsletter *Kumbuka*. Publication in east Africa is impossible because as Wandibba (1990, p. 47) observes ‘… the main publishing houses… are all foreign, and their chief interest is in publishing school textbooks…not more specialised books…’ School textbooks are popular with the publishers because they sell quickly and in large volumes. This is encouraged by the educational systems in east Africa which are acutely pyramidal, primary and secondary schools being at the bottom in that order, tertiary colleges in the middle, and universities at the top. These systems clearly militate against the success of a publishing house which would aim at producing books for the post-secondary school students and the professionals. Thus, any private publisher who wishes to venture into the latter market must strike a delicate balance between publishing school textbooks and specialist books.

For archaeology the situation is complicated by several additional factors which have been discussed by Wandibba (1990) at length. Suffice, therefore, to observe that there is only a handful of indigenous archaeologists in east Africa. These, together with other resident archaeologists plus archaeology students at the universities, can hardly constitute a viable market for archaeological publications. Therefore, it is little wonder that even when they are available, such publications are expensive. The cheapest archaeology book, for instance, costs about four hundred and fifty shillings in Kenya ($20), that is, about 5% and 3% of a middle grade lecturer’s and a professor’s basic monthly salaries respectively, before tax. This is notwithstanding the diminished values of incomes in the country. If such a book were to be ordered direct from an overseas bookseller, the cost would be much higher because of the depreciation of the local currency against hard currencies and because of the packing and postage charges. The situation is worse in Uganda and Tanzania.

The Kenyan situation, however, is ameliorated by the personal libraries of the few local archaeologists. Each of them has had part of their training overseas and, therefore, had an opportunity to buy some publications on archaeological research in east Africa. But, as one would expect, these acquisitions are meant to serve the research interests of the archaeologists concerned. More important, however, are the National Museums' and the British Institute's libraries which have a fairly varied collection of archaeological publications between them. Unfortunately, these libraries are situated in Nairobi, where all resident Kenyan archaeologists, with the exception of one who is stationed at Mombasa on the Kenyan coast, are to be found. This pattern of centralization is not unique to Kenya; it is evident in Tanzania and worse in Uganda. Naturally, it places severe limitations on dissemination of archaeological information in the region. Thus, while the few archaeologists know each other well and share archaeological information to a significant extent, other east Africans— surprisingly even at university level—remain ignorant of what archaeology is (see Wandibba 1990). This probably explains why archaeologists are generally known as 'grave diggers' in the region!

Wandibba (1990, p. 49) has attributed this ignorance to 'the paucity of coverage that the archaeological profession receives from the mass media which only rarely report on archaeological work.' Moreover, the coverage, when it exists, is restricted to new discoveries of hominid fossils and excavation of recent human remains. Additionally, ignorance could be attributed to monopolization of archaeological information by archaeologists. In fact, one could justifiably accuse archaeologists of talking to themselves since they hardly do anything to popularise their profession.

Archaeological data are usually the product of sampling. This is because archaeological sites are often very large. To give one example, Gogo Falls, a neolithic/iron age site located in western Kenya has an area of about 2,500 m^2. The present author has been engaged in research there since 1989 and has so far managed to investigate only a very tiny portion of the site. To excavate such a site entirely would require vast amounts of money, time and manpower, all of which are generally in short supply. In addition, excavation is a destructive process—archaeologists, unlike oral historians, for example, can only interview their informants once because they kill them in the process. Consequently, modern archaeologists prefer to excavate only part of a site as opposed to the whole of it.

To guard against major loss of information during excavation, archaeologists keep what they regard as proper and detailed documentation of the excavation process. The subsequent analysis also generates a large pile of papers. In short, there is a huge amount of documents to study and eventually to store at the end of every research.

The information contained in these documents is crucial, not only to the interpretation of research findings, which usually consist of very large data sets, but also to any subsequent critical evaluation. Unfortunately, dissemination of this information by way of publication is hardly possible. As a general rule, publication means omission of some of the data and illustrations and compression of the rest in order to reduce the manuscript to a size acceptable to publishers. Naturally, in most, if not all, cases, this means much important information is lost, and the usefulness of such a publication is correspondingly decreased. Surely, we cannot claim to do justice to the research findings resulting from several months or years of excavations and analyses by publishing only one or two papers in a journal. Obviously, such a publication highlights what the author thinks is important insofar as his/her specific research objectives. Consequently, the reader can only evaluate the research and its findings within the framework provided by the author. Books which are published at the end of a given piece of research are no exception.

Moreover, since we all know that archaeology has numerous specializations and that it is virtually impossible to get all the specialized manpower one might need during research, publications do not present a comprehensive picture of the potential richness of the archaeological data. For instance, publications on the east African neolithic and iron age have largely been on artefacts, precluding other important aspects like settlement and subsistence patterns. This notwithstanding, it is generally assumed that neolithic and iron age populations were pastoralists and mixed farmers respectively. Although this assumption is not based on systematic research, nearly every archaeologist writing on these periods states the assumption as if it were a fact (eg. Phillipson 1977; Phillipson 1985).

What appears to be different from this general trend of research for publication is research done for the award of university degrees, in which case, most of the information collected during research is made available to the reader. This is because the immediate concern is not to publish but to produce a thesis. However, the problem arises when a reader of a given thesis needs to consult primary information contained in the vast research documents. Availability of these could be a problem because of the lack of, or poor, curation, or even failure of the researcher to deposit copies at the relevant institution. Even when available, the author's experience in east African museums has shown that going through the documents is a tedious exercise which, in some cases, is not worth the trouble.

Computer networks

This state of affairs clearly demonstrates that there is need for a more effective and efficient way of disseminating archaeological information in east Africa. Ideally, use of computer networks could improve the situation. For instance, these could be used in documenting information about archaeological research through the stages of planning, execution and eventual publication. This would include information about research strategies, field observations, the actual data and the tools for analysing them. Storage of all this would be in far less bulky form than ordinary paper, thus easing the prevalent problems of document curation and deterioration.

This would enable readers to evaluate the entire research, including its findings and conclusions, with as little bias as possible, since computers are able to retrieve previously stored information fast and accurately. Moreover, computers allow traditional data processing tasks to be performed faster, which not only greatly reduces the amount of time, money and energy spent processing archaeological data manually, but also eliminates the possibility of making errors (Davis 1981). In addition, computers could make it possible to experiment with new and different methods of data processing. They could, for instance, be used to investigate information processing in archaeological systems or processes by use of simulation. This would involve experimentation with a mathematical or quantitative model representing an archaeological system or process such as domestication. This would not only enable us to explore the behaviour of the people concerned, but also to attempt to offer logical explanations of the problem itself (see Hamond 1978; Aldenderfer 1981).

In east Africa computers could have the additional advantage of enabling researchers to store information concerning other archaeological specializations besides their own. This is necessary because more than half of the region’s resident archaeologists are specialised in one aspect of archaeology only. So far, the region has no environmental archaeologist or palaeoethnobotanist. Therefore, since archaeological research cannot be held in abeyance until every required specialist is available, each researcher could collect as much information as possible concerning his/her specialization and those of others and store it in the computer. This would then be made available to interested researchers as and when required. Thus, use of computers in this way could facilitate dissemination of archaeological information. It could also enable archaeologists to compare and contrast their researches in detail and with considerable ease.

However, it should not be construed that computers have no problems. Thus, although they could encourage and indeed democratise the availability of archaeological data, they could also be used to spy and pry or even to destroy other people’s research. Certainly, misuse of this kind could lead to infringement of copyright. Only if proper care is taken to protect users from such callousness will archaeologists be happy with democratization of their data.

However, the basic problem is that in east Africa there are no computer networks available for archaeologists. The issue is not how networks can be expanded (as in the so-called developed countries) but how they can be set up. Naturally, this calls into question the economics of establishing such a network for archaeology in a region with a chronic shortage of hard currencies which would be required for the importation of computers and related accessories and for their maintenance. In short, as of now and in the foreseeable future, computer networks for east African archaeologists are an impossibility.

Conclusions

Archaeology in east Africa has a very important role to play. This is because it is used together with oral traditions and historical linguistics to reconstruct the region’s pre-colonial history which was greatly distorted and impoverished, and even denied, by colonialism. In fact, because it goes beyond the realms of both historical linguistics and oral traditions, it is considered to be a very important source of local history. With regard to the entire world, east African archaeology provides humanity with information about its ancestry and cultural heritage and about how people’s ancestors related to their environment through time. Furthermore, display of the region’s archaeological finds in museums provides entertainment to the general public.

Given the situation outlined in this chapter, it is little wonder that only a few east Africans, notably the few resident archaeologists plus some historians and linguists, know about the importance of the region’s archaeology. This situation certainly calls for improvement. One way in which this could be done is by systematic popularization of archaeological information by all resident archaeologists. For instance, free public lectures and exhibitions of archaeological material could be arranged. These need not be held only in the large urban centres. Rather, they could be taken to where the people are, that is, to the administrative centres and schools in the rural areas. Additionally, local people could be encouraged to participate in archaeological researches done in their neighbourhood. This would, in turn, present archaeologists with an opportunity to educate them in archaeology. Consequently, the media would get involved in disseminating archaeological information because of the realization that there is considerable public interest in it. In short, popularization of archaeology in east Africa is a necessary, though not sufficient, precondition for effective dissemination of archaeological information amongst the public.

As for the professional archaeologists, archaeological information could be disseminated more efficiently and effectively by adopting cheap methods of making their data available. For example, seminars where all resident and even visiting archaeologists can discuss their researches could be arranged. Joint projects could also be encouraged in order to pool limited resources and expertise. In addition, the commendable practice of donating publications and dissertations and theses to national museums in east Africa, by both local and foreign researchers, should be encouraged, since it is the only major way in which relevant publications from overseas are procured. It would also be useful to explore possibilities of producing cheap archaeological publications locally because, as of now, the problem in east Africa is not democratization of archaeological data *per se*, but democratization of the already available information. Indeed, whilst computer networks could alleviate the problem, especially with regard to future researches in the region, they are yet to be made available. Probably, sensitization of the taxpayers and the decision makers (of course) with external material support could help.

References

Aldenderfer, M.S. 1981. Computer simulation for archaeology: an introductory essay. In *Simulation in Archaeology,* Sabloff, J.A. (ed.), 11–49. New York: Academic Press.

Ciochon, R.L. & J.G.Fleagle 1985. *Primate Evolution and Human Origins.* Menlo Park, California: Benjamin.

Davis, G.B. 1981. *Computer Data Processing (Second Edition).* New York: McGraw-Hill.

Gowlett, J.A.J. 1984. *Ascent to Civilisation: the archaeology of early man.* London: William Collins.

Hamond, F.W. 1978. The contribution of simulation to the study of archaeological processes. In *Simulation Studies in Archaeology,* Hodder, I. (ed.), 1–9. Cambridge: Cambridge University Press.

Isaac, G.L. & E.R.McCown (eds) 1976. *Human Origins: Louis Leakey and the east African evidence.* Menlo Park, California: Benjamin.

Johanson, D. & M.Edey 1981. *Lucy: the beginnings of humankind.* New York: Warner Books.

Phillipson, D.W. 1977. *The Later Prehistory of'Eastern and Southern Africa.* London: Heinemann.

Phillipson, D.W. 1985. *African Archaeology.* Cambridge: Cambridge University Press.

Tobias, P.V. (ed.) 1985. *Hominid Evolution: past, present and future.* New York: Alan R.Liss.

Wandibba, S. 1990. Archaeology and education in Kenya. In *The Excluded Past: archaeology in education,* Stone, P. & MacKenzie, R. (eds), 43–49. London: Unwin Hyman.

4

Polish archaeology and computers: an overview

ARKADIUSZ MARCINIAK & WŁODZIMIERZ RCZKOWSKI

Introduction

To say that computers are fundamental tools in Polish academic research, and archaeology in particular, is untrue. Indeed, although computers may be present in the inventory of available research tools this does not necessarily mean that they are utilized effectively. However, the manner in which computers are used in research procedures can reveal the level of consideration that has been given, as well as the level of expertise available, towards making maximum use of their potential. In this chapter we appraise the state of computer applications in *published* Polish archaeological studies and assess their status.

Application areas

The earliest computer-applications to Polish archaeological problems occurred in the mid 1970s. These were intended to provide objective descriptions of archaeological finds to facilitate classification, and were concerned primarily with only the largest archaeological collections, particularly pottery (eg. Buko 1981) and lithics (eg. Sobczyk 1973). Interest in these fields had developed under the influence of D. L.Clarke's (1968) proposals. Unfortunately, this early work did not stimulate many further computer-based investigations, even in projects which involved the classification of vast collections of pottery. This does not mean that statistical methods in archaeology have been neglected (see Szafranski 1986 for instance), but rather reflects the very low availability of hardware.

After the initial attempts to apply computer methods to Polish archaeology, progress came to a standstill. The situation was changed by an important article by Urbańczyk (1980) on computer methods in archaeology, which represented the first attempt to make some of the more important results of American and British experiences in archaeological data processing and analysis available to the Polish reader. This chapter presented the ability of the computer to verify, sort, modify and reduce data, to present it graphically, to analyse it comparatively—including the determination of similarities, numerical taxonomy, and automatic chronological seriation—and to model data using statistical and simulation methods. Although this paper was essentially a review of contemporary computer applications in Anglo-American archaeology, it remains the standard Polish guide to the selection of appropriate methods.

We can discern a gradual increase in interest in computerized applications during the 1980s, as papers began to address problems in the computer processing of primary archaeological data.

Experience had already engendered critical reflection on the cognitive potential of machine-assisted analysis. Researchers realised that formalised descriptions of archaeological material did not necessarily create the conditions for eradicating subjectiveness. This remained an inherent feature of the process of posing questions, of selecting the data for processing and of interpreting the results of the investigations (Urbańczyk 1980, p. 246).

During the 1980s, Polish archaeologists using computers concentrated their efforts on a few fields of application. Particularly notable were the attempts to create databases of different categories of archaeological material. These were intended principally to be registers of the data coming from the 'Field Survey Project' (FSP). The aim of the FSP was to document all available archaeological sites in Poland. The concept behind this project was to develop an idea of R.Mazurowski (Poznań), who had tried to build a program which would be both a database system and a tool for settlement analyses and even for formulating rules of settlement strategies. Needless to say, this work led to a considerable increase in the quantity of material to be processed.

The *pro-forma* record for each site includes such basic data as cultural and chronological description, information on aspects of the geographical environment and suggestions for preservation. This kind of data constitutes the backbone of a database system produced by Andrzej Prinke and Bogdan Gliniecki in the archaeological museum of Poznań on IBM PCs. The system is very flexible, enabling the user to retrieve data in almost any combination and to generate formatted hardcopy output. This system is now an obligatory standard for all local antiquity services possessing the appropriate equipment. Ultimately, it is hoped that this programme will produce uniform, comparable, collections of data for all archaeological sites in Poland. This will help fulfil a variety of aims, both scientific and professional. So far, such databases have been established only on a local scale in relatively few centres, with the work at Poznań being the most advanced.

The power of graphics in presenting data is well known but, as yet, they have not been applied to Polish archaeology. This is despite Urbańczyk (1987) having described the potential of computer graphics for archaeological research, including the creation of field documentation, its storage and retrieval, the correction of errors, increased legibility through the detection of hidden structures and the extrapolation of the results on to uninvestigated areas, as well as the preparation of materials for publication. The exception is a Polish-Norwegian expedition in Søløy, near Tromsø, which has successfully demonstrated methods of storing, retrieving and presenting excavation data with computer graphics (Bertelsen & Urbańczyk 1985).

More interest has been demonstrated in the area of comparative analysis, where most of the published research relates to investigations into lithic materials, such as the differentiation of the complexes of the Swidry culture of the upper palaeolithic period.

Influenced by Anglo-American spatial archaeology (eg. Whallon 1974; Hodder & Orton 1976), attempts were made to apply computer methods to the analysis of spatial arrangements. A package of computer programs (APIS Archaeological Point Intrasite Sampling) was developed to investigate the early medieval settlement at Wyszográd, in the district of Płock (Masovia), with the representative method. This package included programs for:

- Determining the quantity of network points of a chosen site within a selected area.
- Selecting a sample from this network using a straight or stratified random sample.
- Calculating the surface area.
- Testing the choice of the point dislocation and a plane using the Clarke-Evans method.

These programs enable the package to be used in any archaeological situation where appropriate spatial data concerning the site is available (Kobyliński, Zagrodski & Brzeziński 1984). Continuing interest in spatial dislocation analyses resulted in the publication of a paper presenting the American and British achievements in this area (Kobyliński 1987). This discussed univariate analytical methods, including the 'square' and 'nearest neighbour' methods as well as the bivariate, such as the 'square' and 'distance' methods. In order to make univariate spatial analyses, using nearest neighbour methods, more common, a tabulation module for the 'POISSON' program, written in BASIC V.2, was produced to compute the Clarke-Evans statistic, the Pinder correction, the index of the total nearest neighbour distance (TNND), the Skelan-Moore statistic, and the Pislon and Hopkins statistic. POISSON was written for the Commodore microcomputer.

Numerical taxonomy has found few applications in the work of Polish archaeologists. Exceptions are the typological analyses *of Prdnik* knives of the mid-palaeolithic by Sobczyk (1973) and *Tarwonian* tanged points by Dolukhanov, Kozlowski & Kozlowski 1980.

Published works by Polish archaeologists appear to reveal a similar lack of interest in other more refined techniques such as simulation modelling. A rare application of this approach was adopted in the analysis of the distribution of lithics at the mine of Krzemionki Opatowskie, using a program written in BASICBASIC. Simulation studies of the work at this mine are being continued by the State Archaeological Museum in Warsaw. Apart from the latter example, descriptions of simulation studies are confined to a synthesis of state-of-the-art applications found in the literature of Anglo-American archaeologists (Kobyliński & Urbańczyk 1984).

Assessment

This summary overview of published computer applications in Polish archaeological research allows us to formulate a short conclusion. It should be noted, however, that the level of computer utilization presented above is in some ways misleading. The range of interests of Polish archaeologists, particularly those of the younger generation, is in fact much wider than the one which results from the analysis of the published texts. A conference organised by archaeologists from Warsaw in November 1986 clearly revealed this. Papers presented at the conference covered a wide range of subjects, including a game theory for settlement analysis, numerical analysis of geophysical and radiocarbon data, archaeo-astronomical analyses of prehistoric temples, numerical methods for interpreting figurines and various other statistical applications. This has been, to date, the only such meeting—a forum at which several studies from different academic centres were presented. Unfortunately, most of them have never been published.

It is worth mentioning that the most common application of computers in Polish archaeology is the construction of local databases. These are housed in individual centres, such as Warsaw and Krakow, and are designed to meet particular research needs (eg. metallurgical research, prehistoric settlement systems) in a limited area and for special kinds of material. Experiments in building databases of archaeological material in museums and from excavations, including drawings and photographs, have been attempted by the Archaeological Museum of Poznań; most of the software for this work is written in dBASE III+ and dBASE IV.

Most of the major Polish archaeological institutions have IBM PC/XT microcomputers and, more rarely, PC/ATs. Normally, these institutions will have one or two computers, although the Institute of Material Culture of the Polish Academy of Sciences in Warsaw is exceptional in having about nine, including 4 IBM PC/ATs, and has plans to buy three 80386-based machines in the near future.

Conclusion

The relatively slow adoption and exploitation of computer technology in archaeological research in Poland results from a variety of complex factors. We consider the following to be the most important ones:

- The lack of appropriate computer hardware in most archaeological institutions and, consequently, no network to enable the exchange of information.
- Archaeologists are neither prepared nor trained to use computers; they lack knowledge of the computer's possible cognitive potential, partly as a result of out of date course syllabuses.
- The enormous difficulties in publishing which simply block the flow of information. As a result, it takes four or five years to have a text published.

References

Bertelsen, R. & P.Urbańczyk 1985. The Søløy farm mound. In *Excavations and methods of statigraphical analysis,* Tromsø.

Buko, A. 1981. *Wczesnoś rednicwieozna ceramika sandomierska.* Wroclaw.

Clarke, D.L. 1968. *Analytical Archaeology.* London: Methuen.

Dolukhanov, P.M., J.K.Kozlowski & S.K.Kozlowski 1980. *Multivariate Analysis of Upper Paleolithic and Mesolithic Assemblages. Topology and Ecology.* Warsaw-Krakow: Universytet Jagiellonski.

Hodder, I. & C.Orton 1976. *Spatial Analysis in Archaeology.* Cambridge: Cambridge University Press.

Kobyliński, Z. 1987. Podstawowe metody analizy punktowych ukfedow przestrzennych. *Archaeologia Polski* **32(1),** 21–53.

Kobyliński, Z. & P.Urbańczyk 1984. Modelowanie simulacyjne pradziejowch procesów osadniczych Stan badań. *Kwartalnik Historii Kultury Materialnej* **32(1),** 67–94.

Kobyliński, Z., J.Zagrodski & W.Brzeziński 1984. APIS: omupterowy program próbkowania losowego i statstycznej analyizy przestrzennej stanwiska archeoloicznego. *Archeologia Polski,* **29(1),** 41–55.

Sobczyk, K. 1973. Zastosowaniè taksonomii numerycznej do analyzy typologicznej na przykładzie noży prdnickich. *Sprawozdania z posiedzeń Komisji Naukowych PAN—Oddział w Krakowie* **17(2),** 365.

Szafranski, Z.E. 1986. The application of numeric methods in Poland. *Archaeological Computing Newsletter* **6,** 17–21.

Urbańczyk, P. 1980. Metody komputerowe w archeologii. *Archeologia Polski* **24,** 245–315.

Urbańczyk, P. 1987. Przegld możliwości automatyzacji graficznej analizy archeologicznej dokumentacji terenowej. *Archeologia Polski* **32,** 377–394.

Whallon, R. 1974. Spatial analysis of occupation floors. II. The application of nearest neighbour analysis. *American Antiquity* **39(1),** 16–34.

5

Computer archaeology in Hungary

ATTILA SUHAJDA

Introduction

The Carpathian basin has been a home for various groups of peoples since palaeolithic times as a result of its fertile soils and geographical position in central Europe. Many cultures have flourished and passed away throughout the centuries, and these all left some trace behind them. By the middle of the last century an enormous number of archaeological finds filled the shelves of Hungarian museums, and systematic excavation has only added to the amount. Many of the contents of the museum shelves, including several million prehistoric objects, remained unprocessed. An attempt to solve this problem came with the application of computer processing and database concepts in museums.

The beginnings

In the early 1960s, punch cards were used for archaeological data storage. László Vértes, who collected material from a palaeolithic excavation in Vertesszollos, and Jenö Fitz, whose material came from a roman excavation in Tác-Gorsium, both recorded their finds using punch cards. Vértes then went on to use more thorough methods in the documentation and examination of stone tools.

At the same time, András Mócsy, the head of the Archaeology Department in Eötvös Lóránd University (ELTE) put the names of approximately 80,000 roman citizens (all those that were known at that time) onto punch cards. Following this, for a period of ten years no one carried out any further computer assisted work, and by the end of this time the third generation of computers with their much faster information processing arrived, thus rendering punch cards obsolete. The first article dealing with the new machines was published in 1975 and concerned itself mainly with museum and archaeological data processing methods, giving information about special literature. In 1979 a second article appeared, which dealt with a description of systems previously developed in foreign countries. These included GRYPHOS, SELGEM, and GYPSY. Unfortunately, most Hungarian archaeologists showed no interest in these new computer applications.

The economic policy, as well as the politics and general ideology of the 1960s combined to foster a view of the computer as a wasteful, unproductive toy: something that people should not seriously concern themselves with. This was the case despite the fact that the Hungarian J. von Neumann figured prominently amongst the inventors of the modern computer. It was only possible to introduce suitable equipment when it was available from the west, a procedure hampered by the COCOM list of embargoes, part of the ‘Cold War’ apparatus, and the narrow mindedness of the authorities.

However strange it may seem, these factors played a decisive role in halting further advances in computer science and, to an extent, some of these attitudes carried on into the period of the PCs, with computers still being viewed as something of a mystery.

A growing awareness of the potential applications of computers within archaeology was largely due to the CIDOC congress which was held in Budapest in 1981. One of the aims of the congress was to examine the possibility of Hungary becoming part of a proposed International Museum Computer Network. Unfortunately, the situation was not encouraging, as Hungary was not only lacking the appropriate software, but it did not even have the required computers. However, the meeting was beneficial in the sense that it helped to move the introduction of computers onto a more realistic stage.

In 1982, the Archaeological Systems Improvement Committee was established to ensure uniformity in the description and recording of finds. Some of the leading archaeologists in Hungary came together on this committee with the task of producing a documentation method for finds recording which would be able to satisfy the most diverse requirements. The first aim of the committee was to adopt a database system. Great assistance was given to this end by the Museum Documentation Association (MDA) in the UK and their Museum Documentation System (MDS) was translated into Hungarian.

A description of the MDS system was circulated amongst local museums in the hope of prompting suggestions and advice which might have helped to improve its adaptation to local conditions. Unfortunately, the entire project met with disinterest.

Thanks to this, it was only in 1987 that the new Object Cards produced on the basis of the MDS system came into general use.

The second problem remained: the compilation of a uniform terminology acceptable to everyone, and which dealt systematically with all the various periods. Already in the 1870s the Hungarian archaeologist Flóris Rómer had pointed out the need for such a uniform terminology, but for the partial realization of this issue we had to wait more than 100 years.

An article presenting a concrete plan for progress was doomed to failure as a result of a complete absence of the necessary funding (Halassy 1984). In 1986 there was another turning point. The National Scientific Research Foundation (OKTA), the Hungarian National Museum (HNM), the Archaeological Institute of the Hungarian Academy of Sciences (AIHAS), and the ELTE Department of Archaeology were able to provide some funding for the development of computer and database methods in archaeology.

Equipment and problems

The most widely used PCs in Hungary are the IBM and compatible XT and AT machines. There is a large selection of compatible hardware and software available, produced by a number of firms within Hungary. Unfortunately for archaeologists and for museums, these are the only machines which are within reach of the available budgets, and there are far from enough of even these in operation. The value of computers is still not generally realised and so the demand for them is not great. Larger machines such as those in the IBM System 370 range or a smaller Micro VAX 750 or 780, although found in Hungary, can only be dreamt about by Hungarian archaeologists. The biggest problem is the lack of money for purchasing equipment. It is impossible to produce serious museum record systems or plug into an international computer network without large machines. The arrival of such machines is not foreseeable, at least not in the immediate future.

Another concern is the lack of a uniform attitude which can only be resolved by the creation of the terminology discussed above, and which is a matter of consensus among archaeologists. This situation is compounded by the lack of co-ordinated work between the various institutes and museums. However, this is a problem which can perhaps be more easily solved.

Progress so far

Until now, most of the data processing work has usually been carried out by individuals, but larger bodies of work have also been undertaken, such as the development of a uniform *period* terminology. A large amount of typological description has also been carried out although this, to a large extent, remains inexplicably unfinished.

The first work on computer archaeology is associated with László Bartosiewicz, who in 1981 began a biometrical analysis of animal bones from excavations by the AIHAS using an IBM 3031 computer. A database system in COBOL, BONUS, was produced for the handling of bone data (Bartosiewicz & Bökönyi 1983). This has become of real use since the computers of the AIHAS and the museums were linked into the National Computer Network which was completed last year. Although the National Network IIF (Information Infrastructure Development Program) has now been successfully plugged into the international network, enabling access to information from abroad, it must be said that much work remains to be completed before the structure of the museums' and institutes' systems are in a position to store and adequately ensure a useful exchange of information in this international network.

The exception to this is the HNM, where Bezeczky Tamás has finished the development of MIDAS, a finds recording system suited to local standards. The extent to which other museums will see MIDAS as a useful tool in their work has yet to be seen. Bezeczky was one of the first experts to recognise the possibilities of computers and attempted to publicise their potential. In 1987 he completed a classification system for roman amphorae using the BMDP and CLUSTAN statistical cluster analysis programs on an IBM 3031. As a result, even amphorae of unknown origin can be classified and identified (Bezeczky 1987). Already in operation at the HNM is a graphics program which can be used to store the profiles of ceramic vessels, although only the more symmetrical pieces can be stored in three-dimensions. On the basis of these stored profiles, a classification system for the various classes of ceramics could be developed, probably using cluster analysis, but at this stage nothing is complete. The hardware used comprises one IBM AT, a plotter, a two-dimensional digitiser, and a laser printer. Finds from some of the larger copper age cemeteries are also being processed employing the Bonn Archaeological Seriation and Clustering Package.

Katalin Biró (AIHAS) is at present working on the production of a database for the recording of hundreds of thousands of Central European stone age tools, using an IBM AT compatible and the dBase III+ language. Using this system we can analyse information like material analyses and access specialised literature. LITOTEKA can be used to find all comparative objects outside Hungary, so these areas of the raw materials used for the manufacture of stone tools can be established.

Vera Lányi, working in ELTE AIHAS, deals mainly with Pannonian roman coins. Using an IBM XT and a BASIC database program, she has devised a recording system for approximately 16,000 coins collected from Fejer county, and it currently contains information on all finds made up until 1986. Each record consists of 17 separate pieces of information, six of which are readily available for general publication (Lányi 1989). Ferenc Redö, from AIHAS, is working in a similar field. He has examined the distribution and frequency of coins from the various Pannonian towns using a frequency analysis of time versus quantity (Redö 1986).

At the Velem iron age excavation, Erzsébet Marton is currently working on a more complex database system. Only the most basic information about an object (dimensions, fabric, material, etc.) is stored and there is provision for the grouping of objects with the records of other objects found in a similar locality. The present author is writing a database-assisted program using Turbo Pascal which can locate three-dimensional finds and groups of finds, using this information to identify and draw the various layers of an excavation.

Computers have also been used in the recording of topographic material. There is also a growing interest in this subject in country museums. In the Archaeological Topography of Hungary project, the field survey reports were transferred to a computer database along with the collection of bibliographical references to the sites. In parallel with this work, a team headed by Judith Kvassay is collecting the data for the annual archaeological bibliography of Hungary. However, these projects are limited by the capacity of PCs.

In a slightly different area, pictorial methods of recording have also been experimented with by Pál Raczky (ELTE AIHAS) in the processing of material from the Öcsod neolithic excavation. Experts in AIHAS and the Budapest Technical University have digitised stereo photographs taken of pieces of ceramic material, using an IBM PC/AT clone. The pictures were stored as a series of co-ordinates which obviated the large memory demands of storing the whole picture, although the results are rather less exact that the original version and rather more expensive. However, data concerning 200,000 ceramic pieces will nevertheless be stored on the database, although these will lack the support of a pictorial supplement.

Conclusion

It can be seen that, for all the progress so far, computer archaeology in Hungary has a long way to go. In the 1960s, there was already a demand for computer data storage, but this was doomed through the lack of both finance and interest. In some ways, we do not seem to have progressed very far since then, as a complete database system is still far from complete. The birth of MIDAS is a very big achievement. It seems that a great future awaits it, since the National Computer Network is spreading, and soon all museums will be interested in the application of a uniform system such as MIDAS. However, much good research is still suffering from a lack of funds, and many brilliant ideas still remain paper-bound.

More important, however, than the production of a uniform database for application in all the different museums is the development of specific systems which respond to the needs of various different collections, housed in museums around the country, and yet retain a set of common standards (eg. UNESCO standards). It will be necessary to create a professional advisory body to deal with the technical organization and problems of such a system, as well as to assist in the popularising of the computer.

If a system like the one described above were produced, perhaps even the sceptics would realise that the computer must have its place in archaeological research.

References

Bartosiewicz, L. & S.Bökönyi 1983. Testing the utility of quantitative methods in sex determination of hen bones. *Zoologische Anzeillungen* **210,** 204–212.

Bezeczky, T. 1987. *Roman amphorae from the amber route in western Pannonia.* Oxford: British Archaeological Reports.

Halassy, B. 1984. Basic conditions and possibilities for the development of the Hungarian archaeological database. *Archeológiai értesitö* **111(2),** 252–256.

Lányi, V. 1989. *Fundmunzen der Romische Zeit in Ungarn.* Bonn: Habelt Verlag.

Redö, F. 1986. Münzen als Hilfsmittel bei der Datierung. *Numizmatikai Közlöny* **85,** 34–44.

6
Japanese archaeological site databases and data visualization
AKIFUMI OIKAWA

Introduction

In recent years an increasing number of Japanese archaeologists have come to rely on computers. More than half of all professional archaeologists in Japan use personal computers in their research (if the use of a word processor is included) and several universities and research institutions in archaeology possess large computers. Today, computer applications in Japanese archaeology are sufficiently dynamic to warrant regular symposia to discuss developments.

Computers are primarily used for the following four purposes:

1. The creation and management of databases, which is the focus of this paper.
2. Image processing.
3. Drawing artefacts and features.
4. Statistical analyses.

The management of Japanese archaeological databases

Database management is one of the most important applications of computers in archaeology. In Japan, archaeologists have recognized the potential of computerized databases for organizing and accessing enormous amounts of previously unpublished archaeological data, and various databases have already been created for particular purposes and needs.

There are many reasons for this accumulation of disorganized and unpublished data, but the prime cause is that numerous archaeological sites are being destroyed in the course of building developments. In 1987 more than 20,000 archaeological sites were reported during development and more than 300 million dollars was spent to cover excavation costs. In Japan, 99% of archaeological projects fall into the category of ‘contract archaeology’, and are carried out prior to the destruction of sites. These projects are directed by contract archaeologists, hired by local government, who often have little training in archaeology. Very rarely are professional archaeologists with affiliations to universities or research institutions involved in these projects. Although there are approximately 4,000 contract archaeologists, their numbers are still insufficient to excavate properly all the sites being destroyed. As a result, such contract archaeologists are far too busy excavating to organize and analyse the archaeological data recovered. In other words, finds drawings, feature plans, photographs, and field notes are left disorganised, and data is thus inaccessible to researchers.

Generally, the information stored in these databases is textual and includes details about sites and artefacts, which are relatively easy for computers to handle. Recently, however, some attempts have been made to store image information, such as photographs and scale drawings of artefacts and features. There is a clear need for progress in this field, as excavation yields a considerable amount of image information. In response to this need, the staff of the Nara National Cultural Properties Research Institute are developing a system to store image information in a computer. This includes photographs, slides and rubbings of roof tiles and pottery—all of which resulted from excavations of the Heijo Palace and Capital over the past thirty years. Image databases will be discussed further below.

Visualization

Computers are also used in Japan to compare the morphologies or the surface designs of pottery—formerly a manual operation. Advances in both hardware and software, especially in computer graphics, mean that it is now possible to reconstruct an archaeological site three-dimensionally based on data resulting from an excavation. For example, Ozawa (1988) has developed a solid modelling system which he uses to generate high-quality graphics showing Imperial keyhole-shaped burial mounds and environment. The models are based on measurements taken from individual burial mounds. The database of the keyhole-shaped burial mounds of the Kofun Period (AD 300–600) stores information such as site name, location, date, size, as well as morphological details about the keyhole-shaped burial mounds. We are in the process of developing a program which will allow us to reconstruct the original form of a burial mound based on present-day dimensions.

Management of maps and scale drawings

The use of a computer greatly expedites the mapping of archaeological features at a site and the drawing of artefacts. In Japan, some excavators already rely on a system which exploits a portable computer, hooked up to a laser transit (or electronic distance measurer), to automatically map entire sites or features within a site.

In Japanese archaeology it is considered essential to make precise drawings of an artefact with cross-section and all archaeologically important surface features. This provides an excellent opportunity to observe an artefact and its features carefully, but it can be time-consuming. For example, it sometimes takes a well-trained archaeologist a few days to draw a complete piece ofJomon pottery with intricate surface decorations. However, devices such as the 3SPACE Tracker or the Vectron, enable the same work to be completed in just a few hours. Indeed, laboratories at many local archaeological research centres already utilize a three-dimensional digitiser to draw artefacts.

Other applications

Japanese archaeologists have become very familiar with multivariate statistical analysis. Some specialists apply factor analysis to the results of their measurements of stone tool assemblages for interpretation. Computer simulation has also become a relatively common feature in archaeological research in Japan. For example, Koyama & Sugito (1981) attempted a simulation of population in the prehistoric Jomon Period based on the distribution of sites.

A site database

We have already noted above that individual researchers and institutions have created their own archaeological databases. Unfortunately there is a problem in that recorded data and the structure of databases vary amongst institutions. In order to alleviate this situation, staff and affiliates of the Nara National Cultural Properties Research Institute initiated a project in 1990 to create a common archaeological database format suitable for all researchers and institutions.

As part of this project, we have defined the structure of archaeological data to consist of primary, secondary and reference data. Primary data refers to the actual sites, features and artefacts. Secondary data is derived information, such as scale-drawings of artefacts, photographs, rubbings and so on. Reference data include site reports, theses and administrative documents on contract archaeology, for example. These data sets differ in terms of the contents, where they were produced, and the order in which they were produced; it is impossible to manage just one database for all of them. Of course, within each category (eg. primary data), there exists a hierarchy of relationships (eg. a site yields features which yield artefacts), and the structure of the database must reflect the relationships.

We have decided to set up a national database of sites, for three main reasons: firstly, the great majority of our archaeological data ultimately comes from *sites;* secondly, it is relatively easier to format site databases than artefact or feature databases; thirdly, there are strong administrative and social demands in Japan for information about what kind of site is located where.

At present, we are compiling a site database which, when completed, will contain the following categories of information:

- Identification of sites in Japan (nationwide site database).
- Identification of sites in local administrative units (site database according to city, town, ward, and village).
- Site name (in Chinese characters).
- Site name (in phonetic Japanese characters).
- Location of a site.
- Type of site (such as shell mound, burial mound, etc.).
- Date of site usage.
- Research history of site.
- Map sheet number (the number assigned to a 1/25,000 scale map, prepared by the Geographical Survey Institute of the Ministry of Construction, in which a site appears).
- Coordinate of site (the longitude and latitude of site).
- Synopsis of site (a brief introduction to site).

Japan is divided into 47 prefectures, including the four special ones of Tokyo, Hokkaido, Kyoto and Osaka. All 47 are further divided into cities, towns, wards and villages. Generally, it is these local governments that carry out contract archaeology projects and are responsible for the protection of cultural resources. Consequently, the officials of these local governments are most familiar with local sites.

The actual database will be prepared by local administrative units, such as cities, towns, wards, and villages.

The database will consist of two parts: one whose structure is common to all parts of Japan, and the other whose structure can differ according to local administrative units. Each local administrative unit can create its own site database for inclusion in the second part of the national database. Details of several of the most important Japanese archaeological databases are given in the appendix of this chapter.

In fact, some local governments have already created their site databases. For example, the one for the Saga Prefecture stores information about approximately 8,000 sites, covering almost all the known sites in the prefecture, and the information is readable by a computer. The location of a site is specified by longitude and latitude, and a computer output can show the locations of sites on a map (Oikawa 1985). Once the structure of the nationwide site database is determined, such local site databases will be converted into the nationwide structure.

Another national kind of site database is that of shell mounds (Oikawa & Koyama 1984). It stores the information about approximately 4500 shell midden sites (roughly 95% of such sites in the nation). Since data of all plant and animal remains have been input, it is easy to search, for example, for sites which have yielded both short-necked clams *(Tapes Japonica Deshayes)* and Lamarcks’ hard clams *(meretrix lisoria Roding),* or those which have yielded bones of horses and wild boars. This data is in the final stage of proof-reading and double-checking for accuracy and the next step will be to add the site coordinates to the database. Besides site name, location, and time period, this database contains detailed information about shells, echinoderms, fish, amphibians, reptiles, birds, mammals, human skeletons, and plants.

Image retrieval

All the archaeological databases described above are being used in many ways, but most commonly as a source of information and reference. Recently, there has been a demand to extend database retrieval beyond purely textual queries to facilitate the retrieval of image information, which would enable archaeologists to prepare graphs or two- or three-dimensional images. In response, we have started to develop a system to output two-and three-dimensional images of distribution maps of sites. This system is called VISA (Visualization System for Archaeological databases), and we have finished the basic plan of the system, part of which is already operational. Although VISA at present requires both a host computer which manages databases and a terminal (workstation), we are planning to develop a version of VISA which does not require the use of a host computer.

VISA utilizes several pre-existing software packages: FAIRS (Fujitsu Advanced Information Retrieval System developed by the Fujitsu Corporation of Japan) for indexing and referencing; S (software for data analyses; and graphics developed by the Bell Institute of the American Telephone and Telegram Company) for statistical analyses; and UNIRAS (developed by the UNIRAS Corporation of Denmark) for visualization. Merits of using such pre-existing software include less expensive cost for development, easy maintenance, expansibility, and ease of response to the varied demands of researchers. Researchers can utilize the individual software packages, but they must first be experienced, if not expert, users—which is the major obstacle for any archaeologist who is unfamiliar with computers (the great majority). VISA in this sense acts as a *bridge* between such archaeologists and such software; it is planned so that archaeologists can use this system as long as they know what kind of data sets are to be processed in what way.

The archaeological user enters commands in a question-and-answer format, but the actual data processing is done in the batch format. Its procedure is as follows:

1. Users enter the name of the database in which the information sought is stored (eg. ‘shell middens’) as well as the conditions of the search (ie. what kind of information is sought, such as ‘sites which have yielded short-necked clams’).
2. VISA converts various parameters which have been input to the batch format of FAIRS for the search, thereby transferring the control over the data processing to FAIRS.
3. Results of the search by FAIRS are converted into the input format of UNIRAS and S, and the file in which the results are stored is transferred to a workstation.
4. Control over the data processing may be passed to UNIRAS or S.

Since some researchers require very special diagrams or processing that cannot be covered by UNIRAS or S, VISA allows users to attach and drive their own programs. In some cases, UNIRAS and S can be used from within a program which a user has supplied.

Another feature of VISA is that it incorporates a database of numerical information about the national geography in the system. This database is based on the one, originally created by the Geographical Survey Institute of the Ministry of Construction, which stores the geographical information about Japan in numbers and signs. VISA currently manages information about:

1. coast lines.
2. heights of mountains.
3. classification of topography.
4. lakes, swamps, and rivers.
5. administrative divisions.

Such information is essential for visualization of archaeological data, such as the drawing of a topographical map of a region surrounding a site.

VISA is a system under development; there may well be addition and modification of its functions according to archaeologists' demand. Some possibilities are: processing moving images such as animations, incorporating a simulation program into the system, and expansion of data to be processed (beyond the site database). Further development of this system requires the active participation of archaeologists.

Appendix—important Japanese databases

An important and interesting database is that of the wooden tablets discovered in the Heijo Capital site in Nara. The wooden tablets are thin, irregular strips of wood which were used for writing upon, mainly during the 7th and 8th centuries. They were used as labels, to issue directives, for record keeping, for practising calligraphy or jotting down notes. Some were even used as talismans or title tags. Those used as directives mainly deal with the movement of goods and people within the palace precincts, such as requests for certain commodities, summonses, permits to enter the palace compound, and so forth. Records generally deal with the receipt and disbursement of goods and the work performance of government officials. More than 70% of such wooden tablets have so far been discovered in the Heijo Capital which was the capital of Japan from AD 710 to 784, during the reigns of eight successive emperors. The city was laid out on a square grid pattern modelled on that of the Chinese Tang Dynasty (618–907) capital at Chang'an. Nine major streets running north to south intersected with ten major streets running east to west so as to form 72 large blocks each. Each large block was approximately 550m square. The entire city measured perhaps slightly more than 4.8 by 4.3km.

Another important database is devoted to the roof tiles discovered in the Heijo Capital. This database records the find-spot, type, and date of each roof tile. At present, the Nara National Cultural Properties Research Institute is developing a program that will allow this database to store the images of the rubbings of round eaves tiles (with decorations).

There is also a database of clay figurines ('dogu') of the Jomon Period (c. 10,000–300 BC). This database maintains the information about sites where figurines have been discovered, including such details as the date, type, and morphology of the figurines. It also stores the images of these figurines, which may be output as photographs. Figurines were already being made in eastern Japan as early as the Initial Jomon Period (c. 7500–5000 BC) and had become quite numerous by Middle Jomon times (c. 3500–2000 BC). First made as flat, two-dimensional images, they gradually developed three-dimensional features in the Late Jomon Period (c. 2000–1000 BC). Figurines discovered in Late Jomon sites throughout Japan belong to this latter type.

Saga Prefecture, Hyogo Prefecture, Nara City, Kyoto City, Osaka City and a few other local government departments have created their own site databases.

Database of shell mounds: heaps of shells of kitchen middens left by shell food eaters, dating from the Jomon (c. 10,000–300 BC) and to a lesser extent, the Yayoi (c. 300 BC–AD 300) and Kofan (c. 300–600 AD) Periods. Besides shells, these mounds usually contain other plant and animal food remains, discarded materials, especially pottery, and even burials. Most are found along the eastern coast of Japan where water and natural sandy inlets and beaches provided good conditions for shell fish breeding. Shell mounds assume a variety of shapes and sizes, depending on factors such as their location and length of deposition. Shells may be loosely scattered over a small area or be buried several meters deep on sloping terrain. Shell mounds of the Earliest and Early phases of the Jomon Period (c. 7500–3500 BC) are only of modest dimensions, but in Middle and especially Late Jomon times (c.3500–1000 BC), they increased greatly in size. Dependence on shellfish decreased after the beginning of wet rice agriculture in the Early Yayoi Period (c. 300–c. 100 BC); however, sporadic disposal of shells in midden form has continued in peripheral areas until modern times.

Database of site reports: this is the bibliographic database of archaeological site reports (in Japan more than 3000 site reports are published every year). Besides bibliographic information, this database stores information about the name of a reported site, its date, the type of site, and its location.

References

Koyama, S. & S.Sugito 1981. Jomon Jinko simulation (Simulation of the population during the Jomon period). *Kokuritsu Minzokugaku Hakubutsukan Kenkyu Hokoku (Bulletin of the National Museum of Ethnology)* **9(1),** 1–39.

Oikawa, A. 1985. Site Database for Saga Prefecture Japan. (Unpublished ms.).

Oikawa, A. & S.Koyama 1984. A Jomon shellmound database. *Senri Ethnological Studies* **9,** 187–200.

Ozawa, K. 1988. *Zempokoen Fun no Suri (Mathematics of Keyhole-Shaped Burial Mounds).* Tokyo: Yuzankaku.

7

Archaeological data in the USSR —collection, storage and exploitation: has IT a role?

VICTOR TRIFONOV & PAVEL DOLUKHANOV

Knowledge about Soviet theoretical archaeology outside the country is patchy in the extreme. It was thanks to the enormous efforts of Leo Klein twenty years ago that the outside world came to hear of the very existence of Soviet archaeology. It took longer still for the outside world to become aware of the actual nature of this archaeology and the ways in which it could be 'useful' in practical archaeology, despite the Marxists' branding of Western research as mere 'vulgar materialism' and 'dogmatism'. To judge from current Anglo-American writing in our field it still appears not to be clear to our foreign colleagues (including those who have intimate knowledge of the theory and practice of Marxism and its terminology), what riches Soviet archaeologists have to offer in terms of theory.

The paradox of the situation may lie in the fact that Soviet theoretical archaeologists and their colleagues elsewhere, despite their differing ideological allegiance, have in fact for some time been following a similar model that could be referred to as the 'simplified positivist' model.

As a result, both scientific schools have the same difficulties in explaining their current preoccupations and they are in fact all looking for a way out of the current crisis in more or less the same direction—through re-evaluation of the content, nature and role of culture in the context of socio-industrial development.

A combination of such factors is enough to call into question any assertion that full information on research in the USSR is being made available to those foreign scholars anxious to have access to it. In the current situation there is little to choose between inadequate information and disinformation, which in its turn serves to confuse and finally separate efforts to study properly the subject under investigation. In other words, the discrepancy between what is actually happening in Soviet archaeology and the image that has been created abroad virtually serves to isolate it from the rest of the archaeological world (a fuller treatment of some causes of the apparent isolation of Soviet archaeology is given in Trifonov 1990).

For both European and Asian archaeology it is dangerous to be unable reliably to assess the role and place of the eastern European and central Asian components in the process of the evolution of European culture. Efforts made to overcome such isolation should be joint ones. As far as the Soviet academic world is concerned, efforts to this end involve computerization, first and foremost to record, store and effectively utilize archaeological database. Without such a system Soviet archaeology runs the risk of chaotic disorganization, especially given the intensity with which excavations are proceeding in the USSR.

The main factor making systemization difficult is the present obsolete method of gathering information: the piles of typewritten reports in the disconnected archives of the various Republics scattered all over the country. Such archives, where no proper initiatives are under way to ensure proper distribution of information, become centres of annual pilgrimage for archaeologists, where the only equipment used consists of pencil and paper, since no photocopiers or proper reprographic equipment are available. Data collection becomes a profession in itself and mere possession of information is seen as a major scientific achievement. Limited specialization, all too common among archaeologists, only serves to add to the problem. It is therefore no surprise that foreign researchers are discouraged by the difficulties they encounter when trying to find their bearings in the maze of modern Soviet archaeology. The fact that some succeed is the real surprise.

This situation has arisen despite the fact that Russian and Soviet archaeologists played a pioneering role in establishing the theoretical concepts on which basis the application of IT technology has developed. The eminent Russian archaeologist V.A.Gorodtsov in a number of publications (Gorodtsov 1908; Gorodtsov 1927) put forward principles of typology (numerical taxonomy) which ultimately were adopted in the various systems of computerized classifications.

Gorodtsov was well ahead of his time in setting the formal principles of scientifically based classification of archaeological records. Later on, the typological principles elaborated by Gorodtsov were further developed and successfully applied to various categories of archaeological material by Efimenko (1926) and Artsikhovskii (1928). After a lengthy interval theoretical investigations related to taxonomy were again resumed in the 1960s, including several formalized codes for the description of various categories in the archaeological record (see particularly Marshak 1970; Sher 1965; Kolchin, Marshak & Sher 1970). Such directions continued (Kamenetski, Marshak & Sher 1975) and are still being explored (Bochkarev 1990).

Several computing laboratories in the Soviet Union are attempting systematic processing of archaeological data using locally developed software, especially the laboratories of the Institutes of Archaeology, Academy of Sciences of the USSR, in Moscow and Leningrad; the Institute of Archaeology, Ukrainian Academy of Science in Kiev; the State Hermitage Museum of Leningrad; Kemerovo University in Siberia; the Centre for Archaeological Research of the Georgian Academy of Sciences; and the archaeological institutions of Vilnius, Riga and Tallinn. The main aim of these laboratories is to create local archaeological data-bases. In some cases attempts aimed at chronological ordering of archaeological data have been made.

It is thanks to the initiative of those people who predicted the present situation of masses of undigested data recorded on paper that, approximately 20 years ago, the computerization of archaeological data in the centres of Leningrad, Moscow and Kiev began. However, implementation of this project came to nothing because of the lack of basic prerequisites—computers and programmers—and first and foremost because of passive acceptance by the majority that it would be possible to continue working with outmoded techniques.

By the 1970s, however, the failure to set up a single centralized computerized system had eroded the optimism of those who supported the project and encouraged the scepticism of those who were convinced that the bulldozer is a far more practical innovation than the computer. Theoretical archaeology went on as before, relying on no more than a pencil and paper. Radical changes of attitude have appeared only recently, confirming once more that only crises can give rise to change. At first, individual archaeologists specialising in one particular field were faced by the task of combining their efforts to solve their own particular, but often none the less difficult, problems. Then came the need for agreement on commonly accepted terms to be used in the collection of data. This was followed by the question of how best to achieve rapid and comprehensive processing of data. Many of these problems have now been successfully resolved with the use of computers which various groups of scholars have been acquiring in their efforts to save and store data which they, as individuals, have collected, and which may well have been duplicated elsewhere in the Soviet Union due to inefficient lines of communication.

The most obvious and attractive advantage of computer processing is the amount of time and effort it saves. Only those who imagined that computers could replace brains have found themselves disappointed. Despite the fact that no more than 15 computers are at present being used within archaeology, the effect still promises to be of decisive importance, especially in those branches of archaeology where increased amounts of data are particularly striking (for instance in the USSR steppe regions).

Groups of specialists using computer processing techniques to record materials connected with burials have been set up in Leningrad, Moscow, Kiev, Kishinyov, Krasnodar and Novosibirsk. Bronze articles of the Eurasian bronze age (Professor Chernykh’s group in Moscow), materials from burials of the bronze age in the north Caucasus (Dr. Andreeva’s group in Moscow) and bronze articles of the early iron age of western Siberia and the Baikal Region (Drs. Minaiev and Subbotin, Leningrad) have all been computerized. Programmes for data systemization and quantitative analysis relating to early bronze age ceramics have also been devised. For example, the Centre for Archaeological Research of the Georgian Academy of Sciences (under the coordination of Professor Lordkipanidze) is concentrating on the setting up of several major databases including one on ‘The archaeological sites of Georgia’ (which includes chronology, administrative structures, terrain, altitude, etc.) and another on ‘Greek imports into Georgia’ (Lecheli, *pers. comm.*).

It is important that data bases compiled by different groups are accessible in practice to everyone who is interested in the material concerned. At the same time, such accessibility does not necessarily mean that there will be less competition between groups than before. Unfortunately, only simple systemization of large-scale data has so far occurred and even the trend towards this process is not clearly defined. It is still hardly possible to keep up with data that is currently being collected. These data alone present urgent problems. In general, Soviet archaeologists anxious to pursue these methods are having to lead an intensive but semi-underground existence, keeping out of sight where possible in the academic community.

The need to try and keep under control the data that is currently being accumulated means first and foremost confronting problems connected with the systemization of the chronology of cultures. Being deprived of normal possibilities for publication of material a system has been devised for the storage and presentation of information. Placing, as it does, emphasis on ensuring permanent accessibility to preliminary archaeological information, Soviet archaeology remains an archaeology ‘for internal use only’, artificially isolated from the international scientific community. That is why, in spite of their efforts, foreign colleagues are in most cases unable to interpret satisfactorily the contemporary problems facing Soviet archaeologists. As for most Soviet archaeologists, they have not, until now, felt the need to make their foreign colleagues aware of these problems.

There is no doubt that sooner or later the isolation of Soviet archaeology will be overcome. The guarantee of unity of archaeology as a world science and focus of public interest is provided by its universal, common subject—namely human culture. The question is: what can be done to accelerate this process? Inside the country this means all-out effort to create and develop a computerized system for the storage and processing of archaeological information at least so as to transform the Soviet data base, enabling it to become part of an international data base. In addition, Soviet archaeologists need to become part of the international scientific community.

It is important to allow those who have access to archaeological information, and who still possess the ability to appreciate the essence of the problems facing Soviet archaeologists, to take into account the achievements of archaeology as a world-wide branch of learning. This would facilitate access to more reliable information than can at present be obtained from the leading Soviet scientific published sources. It is important not to rely merely on existing sources and routes of access but rather to develop the shortest route to integration of Soviet archaeology into the international scientific community, thus stimulating Soviet archaeologists to discover themselves.

Acknowledgements

We are grateful to Katharine Judelson for having improved the English of this chapter.

References

Artsikhovskii, A.V. 1928. *K metodike izucheniya serpov/To the methodics of the study of sickles.* Trudy sekcii arkhologii. Moscow: RANION.

Bochkarev, V.S. (ed.) 1990. *Klassificatsia v arkheologii. Terminologicheskiy slovar'-spravochnik/ Classification in archaeology. Terminological reference-book.* Moscow: Nauka.

Efimenko, P.P. 1926. Ryazanskie mogil'niki. Opyt kul'turno-stratigraficheskogo analiza mogil'nikov massovogo tipa/Ryazan cemeteries. An attempt of cultural-stratigraphic analysis of mass-type cemeteries. *Materially po etnograffi* **3(1),** unknown.

Gorodtsov, V.A. 1908. *Pervobytnaya arkhoelogoya / Prehistoric archaeology.* Moscow: Imperial Archaeological Commission.

Gorodtsov, V.A. 1927. *Tipologicheskii metod v arkheologii/Typological method in archaeology.* Moscow: Gosizdad.

Kamenetski, I.S., B.I.Marshak & Y.A.A.Sher 1975. *Analiz arkheologicheskikh istochnikov (Vozmozhnosti fomalizovannogo pogchoda)/ Analysis of archaeological records (Potentialities of the formalised approach).* Moscow: Naouka.

Kolchin, B.A., B.I.Marshak & Y.A.Sher 1970. Arkhologiya i matematika/Archaeology and mathematics. In *Statistiko-kombinatornye metody v arkheologii/Statistical-combinatory methods in archaeology,* Kolchin, B.A. (ed.), Moscow: Nauka.

Marshak 1970. Kod dlya opisaniya keramiki Pedjikenta/Code for the description of the Pedjikent pottery. In *Statistiko-kombinatornye metody v arkheologii/Statistical—combinatory methods in archaeology,* Kolchin, B.A. (ed.), Moscow: Nauka.

Sher, Y.A. 1965. *O sozdanii Kiberneticheskogo fonada arkheologicheskikh istochnikov s avtomaticheskim poiskom informatsii/On the creation of the cybernetic automatic database of the archaeological sources.* Materialy i issledovaniya po arkheologii SSSR. Moscow: Nauka.

Trifonov, V. 1990. Why so little is known about modem soviet archaeology: archaeological data in the USSR—collection, storage and exploitation. *World Archaeological Bulletin* **5,** 77–85.

8

On the importance of high-level communication formats in world archaeology

JOHN WILCOCK

Introduction

Lip-service has been paid to the idea of data standardization in archaeology for a number of years, but there has actually been scant attention to getting working systems of data communication together, and those that have been implemented have been little used. The advantages of data standardization have been promoted at many conferences, and lively discussions have resulted, but the idea that standards should be set up and adhered to has largely fallen on stony ground.

The need for the international communication of archaeological data has been frequently discussed. With the vast amounts of data now being produced, from urban archaeological sites in particular, computers must be used. Of course there are incompatibilities in both computer hardware and software between organizations, but this is a minor difficulty compared with the problem of personal organization. Many suggestions have been published about the use of computers for the storage and communication of data in the various reports of committees that have been examining the future of British archaeology; these include the Report of the Council for British Archaeology Working Party on Archaeological Records (CBA 1975), the *Frere Report* (1975), the *Dimbleby Report* (1978), the *Longworth Report* (1982), the *Cunliffe Report* (1983) and the *Evans Submission* (1983). All have encouraged further use of computers in archaeology, and have exhorted archaeologists to standardize—but with little effect. Rather more threateningly, the *Hart Report* (1985), made to the Science-based Archaeology Committee of the Science and Engineering Research Council (SERC), told British archaeologists to cooperate and standardize by 1990 if they wished to receive future funding by the SERC. The use of computer systems in British archaeology, and no less in world archaeology, has been a piecemeal development made by organizations and individuals lacking coordination; there has inevitably been much reinventing of the wheel. This could be prevented by central planning, which seems unrealistic, or more probably by voluntary acceptance of standards. Such attempts at hardware standardization as have been made (eg. Stewart 1981) have been doomed to failure because every organization, while paying lip-service to the idea of standardization, always wants to use the latest developments in computer hardware, and purchasing is uncoordinated. Recent surveys of computer usage in British archaeology (Richards 1986; Booth, Grant & Richards 1989) revealed that hundreds of different microcomputer types and software packages were in use, without coordination. The most feasible solution is not based on any particular hardware or software, both of which will inevitably date, but is simply the recommendation that the archaeological world generally must agree on a series of high-level *concepts* appropriate to the recording of all types of archaeological data. The entities to be recorded could be text, photographs, artefact drawings, graphics such as site plans and sections, maps, video sequences, and sound sequences. The recording method should be non-sequential and relational links should be available, so that any direction may be taken in browsing through the data. Text should be unstructured (free-format) apart from being tagged by agreed keywords. Transmission of data between incompatible computer systems must clearly be in the lowest common code, which will be ASCII over standard RS232 serial interfaces and modems, and containing no control codes which might cause malfunction of the software in the incompatible target computer. Such data transfer is now commonplace over satellite electronic mail systems, and thus archaeological data transfer should be achievable provided the common high-level concepts can be agreed internationally. These and other arguments have been rehearsed already by Wilcock (1988; 1989b).

Early work on data communication in archaeology

Site data were probably first put into a real computer by Ihm and Gardin in 1959 (Ihm 1961). Since then there has been an ever-growing application of computers to archaeological databanks, including museum cataloguing and research files. Museum workers, who were not necessarily archaeologists, have done most of this work. Their recommended computerized cataloguing systems were proposed in the USA by the Museum Computer Network (Chenhall 1967a; Chenhall 1967b; Ellin 1968; Cowgill 1973) and the Southwestern Archaeological Research Group (Gaines & Most 1982), in France by Gardin *et al.*, and in Britain by the Information Retrieval Group of the Museums Association (IRGMA) since 1965 (Lewis 1965), and more recently by the successor organization, the Museum Documentation Association (MDA) (Stewart 1981; Light & Roberts 1984). A survey of information systems in archaeology has been produced by Martlew (1984). The problems of data recording on the site of an excavation have been addressed by Graham (1976), Wilcock (1978), Powlesland (1983; 1985) and Booth, Brough & Prior (1984). Early transmissions of data over electronic systems were undertaken by Wilcock in 1968 (Wilcock 1969) and by Gaines in the USA.

Some developments in the UK

In Britain many archaeological sites are recorded on the computers of the county-based Sites and Monuments Record, and the National Archaeological Record (part of the National Monuments Record) at Southampton. Many data sets are or have been maintained on computers for specialist purposes. Some examples are bell beakers (Shennan & Wilcock 1975), stone tools (Callow 1976), roman pottery (Young 1980), roman inscriptions (Wilcock 1981), petroglyphs (Walker 1981), tree-ring measurements (Baillie 1982) and radiocarbon dates (Wilcock, Otlet & Walker 1983).

Some developments in other countries of the world

Computer archaeology has developed chiefly in Europe and the USA, and work is known in France, Germany, Belgium, Italy, Spain, Portugal, Poland, Norway, Sweden, Denmark, Holland, the Soviet Union, New Zealand, Colombia and some African countries.

There has been little attempt at standardization, chiefly because use of computers has been opportunistic and dependent on resources—for those even with the finance, and the ability to purchase state-of-the-art microcomputers, there is a split between the use of the incompatible Apple Macintosh and IBM pc ranges of machines. Use of software packages has also been opportunistic, with the result that a tower of babel has developed. This of course is nothing new—such incompatibility was commented upon in the UK and USA on several occasions during the past 20 years, and the situation has not improved in either country.

A few examples will be discussed. Andresen (1988) comments that the reason for the low analytical exploitation of archaeological data in Denmark is the level of Danish theoretical archaeology. Before the use of personal computers, Danish computer archaeology was restricted to a few people in universities; the same experience has been seen in Scandinavia (Bertilsson 1981) and Germany (Gebuhr & Kampffmeyer 1982). There was much criticism of the state of data recording in Denmark for many years, but no action on the recommendations. However, in 1982 the central databank of the Danish archaeological record (DKC) was set up at the National Museum, Copenhagen. Unfortunately the standardization of codes is a ‘political’ problem. A board has been set up to organize the introduction of computers into museums, which has recommended a hardware standard of IBM compatible pc (in 1985), and recommendations on word processing and spreadsheet software.

Madsen (1988), also from Denmark, comments that the central computers run by the Ministry of Education at Århus and Copenhagen were little used by archaeologists, and that IBM compatibles have been used since 1984. There are, however, few implementations of databases. The existing tradition of central registration in Denmark may mean that the central authorities will be able to set up rules for data registration before too many home-grown systems develop. Indeed, central registration of monuments and finds by grid references is already required by law in the Antiquities Act. It is hoped that museums will tend to choose high quality public domain software instead of inferior home-grown products, and in this way some sort of standard will evolve. The KARK project at Århus is using an IBM PS/2 with WORM disk, PostScript printer, plotter, digitiser and scanner. A Husky Hunter field computer is also used. The software in use for this project is dBase III+, C, FORTRAN, Pascal and Microsoft Windows development kit. Other archaeological software will be imported from UK, Holland and Germany—there is no intention of reinventing the wheel.

Djindjian (1989; 1990a; 1990b) includes not only a survey of the contribution of the French school of data analysis (mostly statistics, eg. multivariate data analysis, cluster analysis and correspondence analysis) but also the contributions of workers from Belgium (Cahen, Martin, Slachmuylder, Gob), Italy (Bietti, Moscati), Germany (Bergmann, Hahn, Stehli, Zimmermann, Scollar), Spain (De Quiras), the Soviet Union (Dolukhanov, Kozlowski) and Poland (Schild, Burdukiewicz). He does not, however, discuss data standards as such.

Arroyo-Bishop (1989) describes the ArchéoDATA system. He comments that past experience has shown that large scale centralised projects have usually failed because they have been imposed from the top down, with little or no consultation, using specialised personnel, and they made unreasonable demands on the archaeologists who were intended to use them. The objective with the ArchéoDATA system has not been to change the present recording system of the archaeologists, but to restructure and standardize the data. He proposes the use of standard location codes based on international telephone and postal codes. Use of the optical disk has been explored, with the proposal that not only new excavation records could be placed on them, but old reports as well; but because of the multiplicity of formats currently available, the project is waiting to see which hardware and software standards will finally be adopted. A good recording system can only be acceptable if it foresees and admits that one day a better system will emerge. The most important thing is the data, and it must be possible to pass this on easily and economically to future generations.

Zarzosa & Arroyo-Bishop (1989) describe the computerization of excavations in Spain, Portugal and Andorra. These projects are proceeding at a high rate but in an isolated manner, with only short to medium term views of the exploitation of the data. In France the work of the centres at Valbonne, Lattes and Tours is described. The computerized national survey programme of the Sous-Direction de l'Archéologie of the Ministry of Culture is mentioned, but said to be slow in implementation on computers on a day-to-day basis. In Spain, archaeology has been devolved to the 17 autonomous regions, with the result that the ability to launch national initiatives in data standardization has been severely reduced. In Andorra the small size of the country has led to rapid computerization, but using IBM-compatible equipment, contrary to the wishes of the archaeologists who preferred Apple Macintosh equipment. Computer archaeology in Portugal is said to be in its early stages.

Scollar (1989), very well known for his developments in geophysical survey and computer archaeology, recommends the use of the IUGG 1980 ellipsoid system of mapping for the future calculation of latitude and longitude values for archaeological locational data.

Some future considerations

The ideal use of computers for recording and examining archaeological data would clearly be as a distributed database, with interactive use of networks and multiple processors (Moffett 1984), allowing incompatible hardware to be connected together. The Oracle system, which enables the same software to be run on microcomputers, minicomputers and mainframe computers, using the SQL standard, is recommended. The windows type operating system (eg. Microsoft Windows 3) and the use of menus, hypertext, computer graphics, videodisks and the like will provide a suitable environment for this in the future (see for example Wilcock 1974; Martlew 1988b; Martlew 1988a; Ryan 1988; Goodson 1988; Ruggles 1988; Martlew 1989; Rahtz, Carr & Hall 1989). For graphical construction of diagrams the well-established Autocad system is recommended. It is of note that these are all *software* standards which are being recommended, which will be unaffected by obsolescent hardware. What is needed is a series of *data* 'high-level' standards to go with them. The data is all important, and it must be 'future-proofed', so that it can be used by future generations of archaeologists.

The example of the Museum Documentation Association

From the middle 1960s the major effort in museum information communication in Britain was undertaken by members of the Information Retrieval Group of the Museums Association (IRGMA), which was renamed the Museums Documentation Association (MDA) in 1977. The draft proposals for an interdisciplinary museum cataloguing system were published by IRGMA in 1969. By the early 1970s pilot studies had been carried out in several museum disciplines, and it was known that the preparation of data for computer input added only about one twentieth of the cost to the existing cost of traditional cataloguing. Nevertheless, IRGMA decided in 1971 that a national computer archive on a central computer was unrealistic. Distributed processing is now proving more equal to the task. Having decided to decentralise, therefore, IRGMA was forced to propose communication between individual local computer systems by some hardware-independent means. Thus the *high-level communication format* proposed by IRGMA in 1969 was born. This consisted of a defined series of categories of information, labelled by keywords. Ensuring direct communication between n2 local incompatible computers would require the writing of $n(n-1)$ programs, for each of the computers would have to communicate with every other one, and this number is unacceptably high. The idea was to use the intermediary of the high-level format; each computer would then require two programs to communicate there and back from the high-level format, and the total number of programs required would only be $2n$. Several such high-level formats have been designed for use in museums, eg. for fine art, geological specimens, history artefacts, mineralogical specimens, natural history specimens, scientific instruments, uniforms, bones and general *museum objects*. They were designed by teams of museum curators to meet museum cataloguing requirements, and all were subjected to a lengthy period of field trials, followed by assessment and redesign, before publication with extensive manuals and computer software.

An archaeology high-level format was similarly designed, but it has not been generally accepted by field archaeologists. The reason for this is that either archaeologists did not want the intrusion of the computer into their work, or they could not afford computers, or they were not willing to discipline themselves to adhere to standards, or were not sufficiently organised in their thoughts to use a structured database system, or a combination of several of these reasons. This chapter is attempting to show that mutual benefits can flow from the efficient communication of archaeological data by computer; but this presupposes that archaeologists are willing to discipline themselves to use such a system, and to adhere to some nominal standards.

A case study

International implementation of the international radiocarbon database commission high-level exchange format

A high-level communication format based on the requirements for the Radiocarbon journal publication format, and on local requirements for results processing at the Isotope Measurements Laboratory, AEA Technology, Harwell, was first proposed at the 23rd International Symposium on Archaeometry at Naples in 1983 (Wilcock, Otlet & Walker 1983). The 12th International Radiocarbon Conference at Trondheim in 1985 endorsed the establishment of a Radiocarbon Data Base Commission, to report in one year on the agreed design of a minimum data entry common to all 14C laboratory data bases. The proposed data categories and suitable keywords were published in the proceedings of this conference (Wilcock, Otlet, Walker, Charlesworth & Drodge 1986), together with examples of retrieval in the international Radiocarbon format and a keyword-organised local format. The first steps towards full implementation were left to national agreements. The UK 14C laboratories met in London in April 1986, and a similar meeting was held in the USA. The details of the UK agreement, with listed categories, keywords and sample printouts in ASCII transmission format and expanded high-level keyword format, have been published by Wilcock (1989a). The British implementation uses the Oracle system, and runs on DEC VAX computers. The database is held at Staffordshire Polytechnic, allowing recording of all 14C results for whatever discipline in any British laboratory, and interactive query of the database. In the USA the implementation is on IBM pc-compatible computers, and allows interactive query on a range of keywords. It has thus been demonstrated that standardization of data categories can be achieved if the will is there.

Conclusion

The feasibility of international agreement on data categories for routine communication of data has been demonstrated for one restricted archaeological field, radiocarbon dates for archaeological purposes. There would seem to be no reason, therefore, why international agreement should not be achieved for other archaeological fields, given the will to succeed. Do archaeologists really want standards? Archaeologists continue to demonstrate, at least in the short term, that they are too busy with their rescue work in some instances to consider the matter. But in the medium and long term the advantage of data standards is clearly demonstrated, as those who have tried to read or transfer an obsolete computer file of archaeological data will know. The impetus for standardization will probably arise from increased use of personal interactive computers, improved communications and the availability of multi-media devices. An exciting future awaits in the recording and retrieval of archaeological data.

References

Andresen, J. 1988. The status of the use of computer Applications in Danish archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 409–416. Oxford: British Archaeological Reports.

Arroyo-Bishop, D. 1989. The ArchéoDATA system. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 69–86. Oxford: British Archaeological Reports.

Baillie, M.G.L. 1982. *Tree-Ring Dating and Archaeology.* London: Croom Helm.

Bertilsson, U. 1981. *Dokumentation-forskning. En faltarkeologisk programdiskussion.* Stockholm: Riksantik-varieambedet.

Booth, B., R.Brough & F.Prior 1984. The flexible storage of site data: a microcomputer application. *Journal of Archaeological Science* **11,** 81–89.

Booth, B.K.W., S.A.V.Grant & J.D.Richards (eds) 1989. *Computer Usage in British Archaeology, Second Edition.* IFA Occasional Paper 3. Birmingham: Institute of Field Archaeologists.

Callow, P. 1976. British and French handaxe series. In *Computer Applications in Archaeology 1976,* 33–40. Birmingham: University of Birmingham.

CBA 1975. *Report of the Working Party on Archaeological Records.* London: Council for British Archaeology.

Chenhall, R.G. 1967a. Museum cataloguing by computer. *Computers and the Humanities,* **1,** 240.

Chenhall, R.G. 1967b. The description of archaeological data in computer language. *American Antiquity* **32,** 161–167.

Cowgill, G. 1973. Teotihuacan Data Bank. *Newsletter of Computer Archaeology* **8(4),** 6.

Cunliffe, B. (ed.) 1983. *The Publication of Archaeological Excavations: report of a joint working party of the Council for British Archaeology and the Department of the Environment.* London: Council for British Archaeology.

Dimbleby, G. (ed.) 1978. *The Scientific Treatment of Materials from Rescue Excavations.* London: Ancient Monuments Board, Department of the Environment.

Djindjian, F. 1989. Fifteen years of contributions of the French School of Data Analysis to Quantitative Archaeology. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 193–204. Oxford: British Archaeological Reports.

Djindjian, F. 1990a. Data analysis in archaeology. *Science and Archaeology* **32,** 57–62.

Djindjian, F. 1990b. A select bibliography of French data analysis applications in archaeology. *Science and Archaeology* **32,** 63–68.

Ellin, E. 1968. An international survey of museum computer activity. *Computers and the Humanities* **3(2),** 65–86.

Evans, J. 1983. *Submission to the Historic Building and Monuments Commission from the Council for British Archaeology Working Party on the Organisation of British Archaeology.* London: Council for British Archaeology.

Frere, S. (ed.) 1975. *Principles of Publication in Rescue Archaeology: report of a working party of the committee for rescue archaeology, ancient monuments board for England.* London: Department of the Environment.

Gaines, S.W. & R.Most 1982. Computerized data sharingv: the SARG example of cooperative research. In *Computer Applications in Archaeology 1982,* 19–32. Birmingham: University of Birmingham.

Gebuhr, M. & U.Kampffmeyer 1982. Uberlegungen zum Einsatz von Kleinrechnern in der Ur- und Fruhgeschichtforschung. *Acta Praehistorica et Archaeologica* **11/12,** 3–20.

Goodson, K. 1988. PostScript for archaeological drawings. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 335–352. Oxford: British Archaeological Reports.

Graham, I.D.G. 1976. Intelligent terminals for excavation recording. In *Computer Applications in Archaeology 1976,* 48–51. Birmingham: University of Birmingham.

Hart, M. (ed.) 1985. *Review of Science-Based Archaeology.* Swindon: Science and Engineering Research Council.

Ihm, P. 1961. Classification automatique des objets de l'age de bronze. In *Comptes-Rendus du Seminaire sur les Modeles Mathematiques dans les Sciences Sociales 1960/1,* 28–33. Paris: Ecole Pratique de Hautes Etudes.

Lewis, G.D. 1965. Obtaining information from museum collections and thoughts on a national museum index. *Museums Journal* **65,** 12–22.

Light, R.B. & D.A.Roberts (eds) 1984. *Microcomputers in Museums.* Duxford: Museum Documentation Association.

Longworth, I. (ed.) 1982. *Selection and Retention of Environmental and Artefactual Material from excavations.* London: British Museum.

Madsen, T. 1988. Prospects for the use of the computer in Danish Archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. (ed.), 417–422. Oxford: British Archaeological Reports .

Martlew, R. (ed.) 1984. *Information Systems in Archaeology.* Gloucester: Alan Sutton.

Martlew, R. 1988a. New technology in archaeological education and training. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 499–504. Oxford: British Archaeological Reports.

Martlew, R. 1988b. Optical disc storage: another can of worms? In *Computer and quantitative methods in archaeology 1987,* Ruggles, C.L.N. & Rahtz, S. (eds), 265–268. Oxford: British Archaeological Reports.

Martlew, R. 1989. Pictures and pedagogy: teaching archaeology using image archives on videodisc. *Science and Archaeology* **31,** 12–19.

Moffett, J.C. 1984. An archaeological network database management system. In *Computer Applications in Archaeology 1984,* 93–100. Birmingham: University of Birmingham.

Powlesland, D. 1983. Pots, pits and portables. *Practical Computing June,* 144–146.

Powlesland, D. 1985. *Archaeology and Computers.* London: Council for British Archaeology.

Rahtz, S., L.Carr & W.Hall 1989. New designs for archaeological reports. *Science and Archaeology* **31,** 20–34.

Richards, J.D. (ed.) 1986. *Computer usage in British archaeology.* Occasional Paper 1. Birmingham: Institute of Field Archaeologists.

Ruggles, C. 1988. Software for the Leicester Interactive Videodisc project. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 531–542. Oxford: British Archaeological Reports.

Ryan, N.S. 1988. Bibliography of computer Applications and quantitative methods. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 1–30. Oxford: British Archaeological Reports.

Scollar, I. 1989. Geodetic problems in archaeological data bases at and within the boundaries of some countries. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 251–274. Oxford: British Archaeological Reports.

Shennan, S.J. & J.D.Wilcock 1975. Shape and style variation in central German Bell Beakers: a computer-assisted study. *Science and Archaeology* **15,** 17–31.

Stewart, J.D. (ed.) 1981. *Microcomputers in Archaeology.* Duxford: Museum Documentation Association.

Walker, M.J. 1981. Petroglyphs. In *Data Bank Applications in Archaeology,* Gaines, S.W. (ed.), 112–115. Arizona: University of Arizona Press.

Wilcock, J.D. 1969. Computers and Camelot: South Cadbury, an exercise in computer archaeology. *Spectrum: British Science News* **60,** 7–9.

Wilcock, J.D. 1974. The facilities of the PLUTARCH system. *Science and Archaeology,* **11,** 16–24.

Wilcock, J.D. 1978. The automated archaeologist. In *Computer Applications in Archaeology 1981,* 49–52. Birmingham: University of Birmingham.

Wilcock, J.D. 1981. STRATA—the microcomputer version. In *Computer Applications in Archaeology 1981,* 112–114. Birmingham: University of Birmingham.

Wilcock, J.D. 1988. Past and future directions in archaeological computing. In *Science and Archaeology, Glasgow 1987: proceedings of a conference on the application of scientific techniques to archaeology, Glasgow, September 1987,* Slater, E.A. & J.O.Tate (eds), 497–507. Oxford: British Archaeological Reports.

Wilcock, J.D. 1989a. The international radiocarbon database commission high-level exchange format . In *Archaeometry, Proceedings of the 25th International Symposium Athens 19–23 May 1986,* Maniatis, Y. (ed.), 91–97. Athens: Elsevier.

Wilcock, J.D. 1989b. Review paper: over thirty years' application of the computer in archaeology. *Science and Archaeology* **31,** 35–44.

Wilcock, J.D., R.L.Otlet & A.J.Walker 1983. *Proposals for a high-level international record structure for radiocarbon databases.* Naples: Society for Archaeometry.

Wilcock, J.D., R.L.Otlet, A.J.Walker, S.A.Charlesworth & J.Drodge 1986. Establishment of a working data base for the international exchange for 14C data using universal transfer formats. *Radiocarbon* **28(2A),** 781–787.

Young, C.J. (ed.) 1980. *Guidelines for the Processing and Publication of Roman Pottery from Excavations.* Occasional Paper 4. London: Ancient Monuments and Historic Buildings.

Zarzosa, M.T.L. & D.Arroyo-Bishop 1989. New directions for archaeological computing in south-western Europe (France, Spain and Portugal). In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 319–326. Oxford: British Archaeological Reports.

Visualization

9

Visualizing spatial data: the importance of Geographic Information Systems

GARY LOCK & TREVOR HARRIS

The spatial perspective in archaeology

Archaeologists have extensive spatial data handling requirements. Archaeological phenomena are underpinned by their unique position in space and time, and by the latent relationships existing between them. It is no surprise that archaeologists have paid close regard to, and often emulated, developments in spatial analysis within geography (Hodder 1977; Bintliff 1986; Grant 1986). Traditionally, geographical information has from necessity been displayed, or visualized, in the form of maps. Archaeologists have made full use of both maps and plans to visualize spatial data at different scales from regional to intra-site. Maps provide an extremely efficient storage medium for spatial data condensing tens of thousands of coordinates and associated attribute data into a single sheet of information. Maps also allow a variety of referencing frameworks including specific location identifiers such as cartesian co-ordinate systems—grid references or site grids—or a variety of other locational or 'nominal' location identifiers such as postcodes, place names or arbitrary identification numbers to be used (Tomlinson 1987; Rhind & Mounsey 1989). Vast quantities of locational and thematic information can be stored on a single map and yet, because the eye is a very effective image processor, visual analysis and information retrieval can be rapid. It is not surprising, therefore, that maps have formed the basis for visualizing spatial data from early historic times.

By spatial information is meant any data which are referenced to one or more points in space. More specifically, the term geographical data is often used when referring to spatial data which relate to phenomena observed on the earth's surface (Tomlinson 1987). This can include information concerned with topography, natural resources, surface geology, land-use, infra-structure, cultural resources, employment, voting patterns, demography, health, wealth and, of course, archaeological data. The handling of such geographic information and the inherent spatial relationships between them, is fundamental to management, planning, and decision-making within the natural and human environments. A recent report by the UK Government into the handling of spatial data, known as the Chorley Report, comments that, 'Most human activity depends on geographic information: on knowing where things are and understanding how they relate to each other' (DoE 1987, p. 7). This does not just apply to present human activity; the recognition and analysis of the spatial organization and usage of past landscapes is fundamental to much archaeological endeavour.

It is of crucial importance at this point to recognise the different types of spatial data, for these are at the heart of problems concerned with manual and computer-based visualization and analysis of spatial information. The four spatial primitives of point, line, polygon and raster data, together with associated attribute data, combine to produce a complete graphical and descriptive representation of spatial data. Point data are described by a single cartesian co-ordinate (as artefact find spot for example). Line or vector data describe a single line of unity (linear ditch or river) and polygonal data describe a closed shape or area enclosing a unified data value (the shape of an archaeological site, or area of a soil type). The raster or pixel may count as a fourth data primitive (eg. satellite imagery) which could be seen as an extension of polygon data but is clearly a distinct representation of a spatial unit. These four spatial primitives not only describe the shape of spatial entities but also their locational position and relationships. Any point, line, polygonal or raster data entity can have attribute data associated with it, resulting in locations having values for any number of descriptive variables. The capability of different methods of spatial analysis to integrate all these types of spatial data within their descriptive and analytical procedures is the focus of this chapter.

There are many advantages to be gained from undertaking map generation and spatial analysis within a computing environment; not the least of which is the ability to link together disparate types of geographic information. Digital mapping systems have been available for some time now although their emphasis has been on the storage and draughting of spatial information rather than on its analysis. This chapter contends that recent technological developments in spatial analysis, in the form of Geographic Information Systems (GIS), have important ramifications for visualizing spatial data in archaeology. The technology of GIS allows the analysis of all types of spatial information within a computing environment. This offers a powerful shift of emphasis towards analysis which facilitates whole new approaches towards map-based data. For the first time it is possible to combine the power of traditional manual map analysis with the power of computer storage and wider analytical capabilities.

Approaches to spatial analysis in archaeology

A basic premise in archaeology is that human activity involves the ordering and use of space. Such activities therefore, are likely to be represented by patterning within the archaeological record. For this reason the spatial component in archaeological analysis is paramount. Since early this century archaeologists have used a variety of techniques to visualize, analyse and interpret spatial patterning within the archaeological record. GIS not only provides a further technique for spatial analysis within archaeology, but necessitates a paradigm shift in the underlying philosophy of working with spatial data.

The map-based approach

An early phase of analysis in archaeology then, which continues to form a major strand in present day spatial analysis, involves the map-based approach. In this basically descriptive approach, items of interest are displayed on maps or plans against a background of other related contextual information. The distribution may take the form of point data while the contextual, or symbolic, information is likely to be a mixture of line and polygonal data elements. It is not unusual for a distribution map to show several environmental attributes such as altitude, drainage systems, soil type and vegetation cover. A map-based approach has traditionally been applied at both the intra-site level in the form of site plans, and at the regional level in the form of distribution maps. In the case of intra-site plans the map might, for example, contain the distribution of pottery types across an excavated area of a settlement along with the display of structural and other excavated features. Similarly, at the larger regional scale, the distribution might involve artefacts or sites on a base map portraying a variety of other thematic and contextual information such as altitude or hydrology. As a basis for interpretation the approach is characterized by an informal, almost intuitive, analysis of the mapped data. This involves comparing, correlating, and commenting on the distributions of features by looking at their overall patterning alone or in conjunction with other ‘background’ contextual information. The emphasis of this approach is on visualizing the contextual bi-variate or multivariate relationships existing between the variables of interest within the total universe of relevant background information. The obvious strength of this manual map-based approach is the ability to visualize and analyse point, line, polygonal and attribute data all at once.

An example of this approach at the intra-site level would be two different species of animal bone found to concentrate in two different areas of a site; or certain types of pottery may appear to concentrate around certain types of excavated feature. Similarly, at the regional scale one type of site may be found to correlate with high ground and another with valley sides or valley bottoms. Such an approach, based on apparent similarities and differences between distributions, has a long history (outlined by Hodder & Orton 1976, pp. 1–10; Clarke 1977, pp. 2–5) stretching back to at least Fox’s pioneering study of *The Personality of Britain* (1932). Inherent within this approach is the recognition that graphical representation is a simple method for the rapid and often crude summary, display and exploration of relationships within large complex spatial data sets. It is this very complexity, however, which makes analysis so difficult. As a result, analysis may only involve an informal cursory search for correlations between a limited combination of archaeological and environmental variables. This is not to say that this approach does not produce results but that it operates within an informal heuristic and intuitive framework. Indeed, the fact that maps are still the main form for visualizing spatial data, sometimes facilitated by the application of graphics or draughting computer software in the map production process, is proof of the overall utility of this approach in archaeology.

The statistical approach

During the 1970s, in common with most social science and humanities disciplines, archaeology was part of the overall movement drawn toward the adoption of quantitative methods for data recording and analysis. These new approaches were concerned with developing more formal and objective ways of isolating and of testing for patterns within data. This, as the work of Clarke (1972a) emphasized, included spatial patterning. Archaeologists began to apply formal quantitative methods of spatial analysis, developed within other disciplines such as plant ecology and geography, to a wide range of archaeological data. These applications were characterized by a progressive emphasis on computing. The publication of Hodder & Orton’s *Spatial Analysis in Archaeology* (1976) brought together, and made available to the general archaeological audience, for the first time the whole range of quantitative techniques then available.

The quantitative approach to archaeological spatial analysis handles archaeological phenomena in the form of either points, made up of x, y co-ordinates, or as counts within the cells of a grid which has been overlaid over the area of interest. The results of these techniques are usually to produce summary statistics indicating a tendency towards clustering or towards regular or random patterning of the archaeological distribution. A statistical comparison of two or more distributions can also be undertaken to indicate the extent to which they correlate spatially. Most tests include some measure of statistical confidence in the results obtained. The challenge facing quantitative archaeologists is to match these formal methods with the ideals of spatial analysis that have been apparent for many years:

> Spatial archaeology deals with human activities at every scale, the traces and artefacts left by them, the physical infrastructure which accommodated them, the environments that they impinged upon and the interaction between all these aspects. (Clarke 1977, p. 9)

It has to be admitted that archaeologists who have employed a quantitative approach to archaeological data analysis have always been in a minority, with the majority preferring the well established informal methods of map-based spatial analysis. This trend exists despite the long recognised problems associated with the informal analysis of maps, including biased samples and the effects of cultural and non-cultural post-depositional alterations to spatial patterns. Even ignoring the fact that many innumerate archaeologists will avoid statistical techniques simply because they are outside their own area of expertise, to the majority these deficiencies must seem preferable to the limitations and problems posed by quantitative methods. Hodder & Orton themselves (1976, chapter 8) recognised that spatial analysis involved firstly testing for spatial patterning and then explaining the cultural processes which produced those patterns. Formal methods of spatial analysis, they acknowledged, could only be applied to the first of these and the relationship between the two is often complex because final patterns may well be the product of several superimposed patterns. The application of an incorrect statistical model could result if a pattern is judged to be random even though it is composed of several overlying non-random patterns which are the results of human behaviour; itself rarely random. Other technical issues also complicate the situation. Results from grid-based techniques, for example, can often be influenced by the size, shape and placement of the cells (Carr 1984, pp. 175–180) and the boundary effects on the nearest-neighbour analysis of point distributions were recognised many years ago (Hodder 1971).

A major restriction with formal quantitative methods is their inability to incorporate contextual, or symbolic, information within the analysis. By this is meant the mass of information that exists in any spatial analysis concerning the sites or artefacts themselves (attribute data) and their physical location on the ground (point, line and polygon data). Statistical methods are severely limited to point data, being unable to integrate line and polygonal entities into the analysis. As mentioned above, at the intra-site level this could include structural evidence for buildings, defences and boundary markers, information about the excavated context of the items such as type of pit, posthole, layer, gully, ditch, its fill or phase, or even descriptive information about the items themselves. At the inter-site or regional level there are numerous environmental data to be considered including topography, hydrology, soil type and vegetation cover, as well as the cultural aspects of the sites or artefacts themselves. It must be acknowledged that a quantitative approach cannot incorporate anything but a small sub-set of the available archaeological data being used. If we accept that the spatial ordering of archaeological data reflects social and cultural behaviour, then to use only a sub-set of that data could result in a distorted interpretation. In terms of trying to visualize spatial data through quantitative methodologies the results can only be described as disappointing. While many statisticians can visualize data in 'n' dimensional space based upon a numerical algorithm and graphically represented in the form of output from a multivariate technique such as factor analysis, principal components analysis or discriminant analysis, most archaeologists are much less statistically sophisticated. To them all that is achieved is the visualization of a set of points in two dimensional cartesian space. Hodder & Orton (1976) acknowledged the importance of contextual information by including a chapter on 'The relationship between sites and other features', although their approach was limited to including only one 'background' variable (eg. soil type or distance from a source). By applying a test of association (such as chi-square) it is possible to comment on the relationship between a point distribution and categories of the background variable. However it is interesting to note that in two important published collections from the same period (Clarke 1977; Hodder 1978) analyses involving contextual information were undertaken but the latter remained external to the techniques themselves.

Probably the most powerful critique of the quantitative approach to spatial analysis appears in the opening sections of Kintigh & Ammerman (1982). By caricaturing the steps of an archaeological spatial analysis they identify the problem areas and limitations of 'traditional' quantitative methods. They suggest an alternative approach of 'heuristic spatial analysis' which aims to 'combine the intellectual sophistication of intuitive approaches with the information processing capacity and systematic benefits of quantitative treatments' (Kintigh & Ammerman 1982, p. 33). The result is still somewhat disappointing in that the quantitative methods developed (based on k-means clustering techniques) are only heuristic in the sense that the methodology is iterative so that new assumptions can be incorporated during each iteration. The whole problem of incorporating contextual information is reduced to a future 'area for development' at the end of the chapter. Further criticisms of standard spatial quantitative methods have been listed by Whallon (1984) who sees the main problems as centring on the basic incongruencies between the statistical techniques and the nature of archaeological data, as well as the nature of the questions asked of the data. This has resulted in 'a large measure of dissatisfaction with available approaches and their results' (Whallon 1984, p. 242).

Just as the variety of data types included in map-based approaches heavily influence the mode of analysis, so too does the limited data suitable for quantitative analysis. *Intrasite Spatial Analysis In Archaeology* (Hietala 1984) is an important collection of papers which illustrate the power and, at the same time, the limitations of the mainstream quantitative techniques. All of the applications use data either in the form of co-ordinates or grid cell counts. Four studies include a limited amount of contextual information (eg. hearths within hunter-gatherer campsites) in their analyses although in each case this is added at a later stage and is not an integral part of the technique. Three other papers incorporate implied contextual knowledge within analyses of building plans where grid cells represent a structural element such as a room.

Many of the techniques described in Hietala (1984) have been integrated into PC-based software[1] (Blankholm 1989). ARCOSPACE has been written specifically for archaeologists and must be regarded as state-of-the-art in archaeological quantitative spatial analysis. The software provides eight spatial analytical techniques all of which require either point or grid data. The inability to include any sort of contextual information, other than the minimal ability to assign items to a 'class' of one categorical variable, within the analyses is apparent. The same restrictions apply to other readily available software; Multi-Response Permutation Procedures (MRPP)[2] (Berry, Mielke & Kvamme 1984; Rodgers 1988) makes no allowance for contextual information other than for categorising points according to type. Attwell and Fletcher's test of association[3] (Attwell & Fletcher 1987) tests for significance of a point distribution across the categories of one 'background' variable, such as altitude, which has been classified into a series of height ranges. New quantitative techniques recently developed are subject to the same limitations. Intrasite Spatial Structure (ISS) (Djindjian 1988) for example, is based on correspondence analysis, a multivariate technique which has become very popular in archaeology over the last few years (Madsen 1988). A recent software package that has attempted to integrate contextual data is SITEPAK[4] (Cribb 1986; Cribb 1987; Cribb & Minnegal 1989). This performs standard point and grid quantitative techniques at the intra-site and regional levels graphically presenting the distributions on background maps of the site or region. While the background information is not directly used in the spatial analysis it is a powerful aid to intuitive interpretation of the complete picture. Excavated features or regional environmental information can be digitized from excavation plans and maps and stored separately to the point and grid data which are used for the quantitative analyses.

The objectives of the quantitative approach, therefore, are firmly embedded within a hypothesis testing confirmatory framework. A given pattern of points are tested to isolate and identify the direction and strength of any relationships between the points. Results are in the form of a summary statistic usually with an associated level of significance. The limitations of these objectives in archaeological terms are exposed by Carr (1984) in a review of quantitative spatial methods. He classifies them (1984, p. 134) according to which of the following three 'archaeological' questions they can be applied: i) are the points randomly scattered, ii) what are the spatial limits of clusters and iii) do clusters overlap. It is interesting to note that many of these quantitative spatial techniques employ graphical representation of the results in an attempt to visualize the data. The importance of the visualization of data is also apparent in the relatively new approaches of Exploratory Data Analysis (EDA, see Tukey 1977). This is just part of a wider trend toward the retaining of visual analysis in statistical work and a move away from the severely reductionist approach of some statistical methods.

While map-based and quantitative approaches to archaeological analysis emphasize the spatial component, the gulf between the two approaches would appear cavernous. Both suffer from severe limitations caused by the character of their respective data requirements: the results of any analysis depend on the quality of the data being used as well as on the quality of the 'statistical' or 'map' model being used. There has always been an undercurrent of cynicism and lack of empathy between practitioners of the two respective approaches caused by the intractable constraints inherent within each. Integration of the two approaches has not been possible because of the inability to integrate the different spatial primitives within the same analytical environment; that of the computer. This is now possible with the technology of GIS. The strengths of such an integrated approach were recognised many years ago by Clarke in his model of the iron age settlement at Glastonbury (Clarke 1972c). This work remains a landmark in spatial analysis in archaeology because it embodies the major strengths of map-based methods yet reflects Clarke's quantitative approach to analysis. This approach permitted the combination of a variety of data types such as artefacts, structures, features and environmental variables, so as to build up as complete a picture as possible. The data were then subjected to 'a developing lattice' of analyses combining vertical spatial relationships, horizontal spatial relationships, structural relationships and artefact relationships. As Clarke states,

> the spatial relationships between the artefacts, other artefacts, site features, other sites, landscape elements and environmental aspects present a formidable matrix of alternative individual categorizations and cross-combinations to be searched for information. (Clarke 1972c, p. 363)

In other words, the more relevant data that can be integrated within an analysis, the wider the range of questions that can be asked thus improving the resulting interpretations. It is this flexibility and power resulting from the integration of all types of spatial data, that are being offered by GIS.

Digital mapping

Having established the fundamental importance of maps to both visualizing and analysing spatial data in archaeology, we must now turn to digital mapping, an area of software development which has been growing considerably for the last decade. It is important to realise the differences between digital mapping/plotting software (CAD) and true GIS (see Cowen 1988; Lindenberg & Fisher 1988). CAD systems such as the PC industry standard Autocad or the mainframe based UNIRAS can be used as graphical front-ends for GIS but they are primarily for electronic drafting or mapping. CAD systems will draft a file of co-ordinate data but are severely limited in the analytical tasks they can perform, for example the integration of spatial entities or the linking of attributes stored in a database to entities stored graphically. While CAD software offers greatly increased power and flexibility over manual mapping, analysis of spatial data is confined to the plotting of different classifications of attributes. The whole emphasis of GIS, on the other hand, is the integration of all types of data within a framework of spatial referencing specifically for the purpose of analysis. It is a common misunderstanding that a database plus digital mapping equals GIS. This is far from the truth. A GIS is an analytical engine that can use digital mapping as just one of several methods of displaying results from a whole range of analytical procedures. In essence, digital mapping systems can not link together discrete data entities recorded as point, line and polygons because the topological relationships between the entities are not recorded. The data may be visualized digitally but it is a hollow shell with no inner substance holding together the individual elements.

To conclude this section we can state that maps have always been an ideal medium for visualizing spatial data, whether they are drawn and stored by hand or, as more recently, digitally. Problems arise, however, when such spatial data are to be analyzed. Not only do manual maps fossilize information because of the severe restrictions in updating information once it has been committed to map form, but they also make it very difficult to perform any analysis that involves data on more than one map covering the same area. The problems are illustrated by McHarg (1969) in his attempts to overlay two maps and trace the areas of intersection or union, an early manual attempt at GIS in fact. This illustrates the severe analytical weaknesses of maps as a method of visualizing spatial data; data integration from non-conforming area locations is not possible. This is a major lacuna given the diversity of spatial data types and sources. Depending on the variables and the analysis there may be good reason to just map the data and not integrate them but in order to gain a critical analytical ability it is essential that a GIS underpins the creation of mapped data.

Geographic Information Systems

Handling geographic information within a computerized system has presented a number of problems. Recent developments in the storage capacity and processing speed of computers, in tandem with significantly reduced hardware costs (by a factor of 100 in the last ten years (DoE 1987, p. 7)), have overcome some of these difficulties. This has stimulated the development and use of systems specifically designed to handle inevitably large quantities of locational information. Furthermore, the ease of use and the reduced cost of ‘off the shelf specialist GIS software has brought such technology within the purview of many non-specialist users. Much geographic information, such as population census data, satellite imagery, utility company network information, and increasingly, topographic information (Mayes 1986; DoE 1987, pp. 59–75), is already available in digital form. Together these developments have facilitated the development of GIS as a tool for undertaking advanced forms of computerized spatial data handling and analysis.

This new technology is anticipated to have a major impact on how geographic information is used. The Chorley Report suggests that the development of such systems is as significant to spatial analysis as:

> the invention of the microscope and telescope were to science, the computer to economics, and the printing press to information dissemination. It is the biggest step forward in the handling of geographic information since the invention of the map. (DoE 1987, p. 8)

Certainly, evidence from North America where many such systems originate and where application areas, including archaeology, are well established (worth an estimated $3 billion annually in 1986 (Rhind & Mounsey 1989, p. 575)) would support such a claim (DoE 1987, pp. 153–161; Abler 1987).

The history of GIS

Despite what appears to be a very recent history GIS have been in existence since the 1960s (see Rhind 1977; Nagy & Wagle 1979; Rhind 1981; Burrough 1986). The Canada Land Inventory, established over twenty years ago, was concerned with converting map data into digital form. By 1972 the Canada Geographic Information System (CGIS) had become the world’s first operational integrated GIS. The system achieved this distinction through its ability to not only input, store and display coverage maps, such as forestry reserves or soil type, but to overlay two or more coverages for a region and compute the areas of simple or compound coverages (Nagy & Wagle 1979, p. 171). In this the CGIS differed from purely cartographic systems and characterizes a primary function of GIS. At around the same time the seminal text on GIS edited by R.F.Tomlinson was published under the aegis of the International Geographical Union (Tomlinson 1972).

Recognition of the potential of GIS for archaeological work first entered the literature in North America in the early 1980s (Brown & Rubin 1982). By 1985 both Commission IV (Data Management and Mathematical Methods in Archaeology) of the Union International des Sciences Pré- et Proto-historiques and the Society for American Archaeology (SAA) had sessions on GIS (Gill & Howes 1985; Kvamme 1985b). Papers covered both methods and principles (Kvamme 1985a; Ferguson 1985) and specific regional applications (Bailey, Howes, Hackenberger & Wherry 1985; Creamer 1985). North America has continued to lead in GIS applications including an account of the principles of GIS, the organization of spatial data and GIS potential in archaeology (Kvamme 1986b) and descriptions of suitable commercial software (Ferguson 1986; Miller 1986). Kvamme (1986a) has developed a specifically archaeological GIS called ARGIS now renamed TERRAIN PAC (see also Limp & Farley 1986) and archaeological applications of leading commercial software such as ARC/INFO are established (Oliver & Schroeder 1986). Overviews are still appearing (Zubrow 1987; Kvamme & Kohler 1988; Kvamme 1989) and a major volume appeared in 1990 (Allen, Green & Zubrow 1990) documenting over twenty archaeological applications of GIS in North America. Awareness of GIS potential has continued to rise in the USA and the situation is still one of expansion although many applications remain as yet unpublished (Kvamme, *pers. comm.*).

The adoption of GIS in archaeology elsewhere in the world has been much slower with the lead coming from geographers rather than archaeologists. In the UK, Harris (1986) showed the need for computerized spatial data handling procedures at the regional level for archival, educational and research purposes as well as for decision-making within the planning environment. He outlined a GIS application to the archaeology of the Brighton area, a theme which was taken up again at the next conference (Harris 1988) with the addition of Digital Terrain Modelling (DTM) and three-dimensional graphics as output from the GIS (see also Harris & Dudley 1984; Harris 1985). Two other geographers have also shown the potential of GIS in archaeology with their work on remote sensing techniques in the Fens (Donoghue & Shennan 1988; Donoghue 1989). The prospects for the widespread adoption of GIS in UK archaeology have been discussed by the current authors (Harris & Lock 1991) as well as the specific application area of cultural resource management (Harris & Lock 1991). Other European GIS applications include the work of Wansleeben (1988) in the Netherlands, Arroyo-Bishop (1991) in France and Stančič in Yugoslavia (Stančič, *pers. comm.*).

The functionality of GIS

Geographic Information Systems are computer-based systems designed to input, store, transform, manipulate, analyze and display spatial data traditionally represented in the form of maps or plans. In essence such systems are characterized by their ability to store many sets of locational data, usually representing a series of map layers, and enable these layers to be compared and integrated. The power of GIS lies in their ability to store not only the locational and attribute data for each spatial entity but also the topological relationships between them. This permits the different spatial features making up each map layer to be integrated with those of other map coverages, examined in the same analysis, and new component maps or information produced.

It is this ability to handle spatially disparate data from several map layers, to seek relationships, to produce composite variables and maps, and to model the information, which makes GIS so potentially important to archaeology. The benefits of visualizing spatial data in the form of maps are greatly enhanced in this way by the underlying ability of GIS to integrate and analyze spatial information and produce new map output. GIS are well suited for ‘what if type queries encouraging an exploratory approach to data analysis. Besides the considerable data management capabilities of GIS, vector data models enable more realistic portrayals of archaeological entities to be recorded for analysis and obviate the crude reduction of archaeological sites to single point data regardless of their actual size and shape. The extensive ‘windowing’ capabilities of GIS also allow considerable flexibility in defining areas of interest anywhere within an apparently seamless map extent.

This is not the place to deal in detail about the full capabilities of these systems and materials elsewhere may be consulted for this purpose (see for example Burrough 1986; Star & Estes 1990; Tomlin 1990). Several features of GIS functionality, however, can be identified which best illustrate the data management, analytical, and graphical capabilities of GIS. A primary feature of GIS is the ability to select, integrate and analyze features from a combination of map coverages and to construct new composite variables or maps from these sources. The ability to overlay map layers in this way enables areas of intersection or union between different data layers to be determined and subsequently used as a basis for further analysis. For example, two map layers portraying soil type and hydrology could be overlaid and relationships sought which could be reclassified to indicate potential agricultural land. This, in turn, could be used as a basis for examining or determining the distribution of site types. Similarly, slope and aspect characteristics can be obtained from altitude data and included in subsequent analyses to establish the preference of certain site types to these criteria.

GIS also possess operations which enable the ‘buffering’ of either point, line or polygonal map features or the generation of separate overlay features. Thus a buffer around a linear feature such as an ancient routeway, for example, would generate a corridor across a computerized landscape within which data could be analyzed or manipulated. The use of such operations for neighbourhood analysis and site catchment analysis of archaeological sites, along with the buffering of environmental or archaeological features which enable the exploration of various hypothesised relationships, provide a very powerful addition to the armoury of techniques available to the archaeologist.

Map data can also be reclassified, and continuous variables reduced to a categorical state. Such an operation permits contingency tables and associated statistics, for example, to be produced to augment graphical information. This ability to transform data, to create new or composite variables from existing coverages which in turn are stored within the database, is a clear indication of how GIS extends the graphical utility of maps to provide extensive analytical abilities.

The flexibility of output formats and the ability to visualize the manipulated map data is a further strength of GIS. GIS is underpinned by traditional methods of digital mapping and graphical display. Output can take the form of two-dimensional maps, three- (or more correctly four-) dimensional surface graphics, statistical and tabular information. Digital Terrain Models, usually based on altitude data, can be produced and geographical attribute data draped onto the surface to graphically portray complex spatial information in the context of landscape form (Harris 1988). Three-dimensional representations of topography with soil type, hydrology, archaeological sites, and other variables of interest, for example, can provide dramatic visual impact of underlying association and, thereby, aid explanation. Questions concerning line of sight, intervisibility and viewsheds can equally be addressed by such information. Statistical and numerical output in the form of descriptive and inferential statistics and more advanced techniques of spatial analysis can also be produced to augment the graphical display capabilities.

Finally, GIS provide an environment in which sophisticated modelling can be undertaken. In this instance archaeological data often suffer from the problem of ‘white areas’ in which no data is known to exist, although it is unsure whether this is due to some bias or represents a true absence of data. GIS offer a powerful modelling environment within which it is possible to generate models to predict and extrapolate beyond the available data and into these ‘white areas’. If, for example, multivariate relationships are known to exist which help explain the location of certain archaeological sites to a high degree, then the ability exists to use these relationships within a GIS to predict and map the location of possible, as yet unknown, sites.

Conclusion

It should be obvious from this overview of GIS functionality that this technology holds the key to future development in visualizing spatial data. The extensive experience of R.F.Tomlinson, a doyen of GIS in North America, regarding the uptake of GIS by prospective users on that continent is particularly revealing,

> in dealing with a relatively new technology such as GIS we have found over and over again in North America that the technical problems are minor in comparison with the human ones. The success or failure of a GIS effort has rarely depended on technical factors, and almost always on institutional or managerial ones…In short we believe that the greatest obstacle to greater GIS use will continue to be the human problem of introducing a new technology which requires not only a new way of doing things, but whose main purpose is to permit the agency to do a host of things which it has not done before, and in many cases does not understand. (Tomlinson 1987, pp. 154, 158)

The last part of the above quotation is worth emphasizing in terms of visualizing spatial data. GIS will not only automate many of the long accepted methodologies used with manual maps but is also capable of whole new approaches to spatial data. The combination of the individual advantages of map-based and some statistical approaches together with the storage and analytical benefits of working within a digital environment, promise an exciting future for GIS in archaeology.

Acknowledgements

Gary Lock gratefully acknowledges the assistance of the British Academy for providing a grant to enable him to attend the World Archaeological Congress 2.

Notes

1 further information about ARCOSPACE is available from H.P.Blankholm, Department of Prehistoric Archaeology, University of Århus, Moesgard, 8270 Hojbjerg, Denmark.

2 MRPPX is a more user-friendly version of MRPP written in FORTRAN 77.

3 Copies of this program (written in GW-BASIC) can be obtained by sending a 5.25 inch floppy disk to M.Fletcher, Department of Mathematics, Staffordshire Polytechnic, Beaconside, Stafford, ST18 0AD, UK.

4 Further information about SITEPAK is available from R.Cribb, Central Land Council, 33 Stuart Highway, PO Box 3321, Alice Springs NT5750, Australia.

References

Abler, R.F. 1987. The National Science Foundation National Centre for Geographic Information and Analysis. *International Journal of Geographical Information Systems* **1**, 303.

Allen, K.S., S.W.Green & E.B.W.Zubrow 1990. *Interpreting Space: GIS and archaeology.* Basingstoke, Hants: Taylor and Francis.

Arroyo-Bishop, D. 1991. The Archéodata system: towards a European archaeological document. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 61–69. Oxford: British Archaeological Reports.

Attwell, M.R. & M.Fletcher 1987. An analytical technique for investigating spatial relationships. *Journal of Archaeological Science* **14**, 1–11.

Bailey, R., D.Howes, S.Hackenberger & D.Wherry 1985. Geographic information processing in land use modeling and testing in the Columbia River basin. Paper presented at the 50th Annual Meeting of the Society for American Archaeology, Denver.

Berry, K.L., P.W.Mielke & K.L.Kvamme 1984. Efficient permutation procedures for analysis of artifact distributions. In *Intrasite Spatial Analysis in Archaeology,* Hietala, H. (ed.), 54–74. Cambridge: Cambridge University Press.

Bintliff, J.L. 1986. Archaeology at the interface: an historical perspective. In *Archaeology at the Interface: Studies in archaeology's relationships with History, Geography, Biology and Physical Science,* Bintliff, J.L. & C.F.Gaffney (eds), 4–31. Oxford: British Archaeological Reports.

Blankholm, H.P. 1989. ARCOSPACE. A package for spatial analysis of archaeological data. *Archaeological Computing Newsletter* **19,** 3.

Brown, P.E. & B.H.Rubin 1982. Patterns of desert resource use: an integrated approach to settlement analysis. In *Granite Reef: a study in desert archaeology,* Brown, P.E. & C.L.Stone (eds), Anthropological Research Papers **28,** 267–305. Arizona: Arizona State University.

Burrough, P.A. 1986. *Principles of Geographical Information Systems for Land Resources Assessment.* Oxford: Clarendon Press.

Carr, C. 1984. The nature of organization of intrasite archaeological records and spatial analytical approaches to their investigation. In *Advances in Archaeological Method and Theory,* Schiffer, M.B. (ed.), **7,** 103–222. New York: Academic Press.

Clarke, D.L. 1972a. Models and paradigms in contemporary archaeology. In *Models in archaeology,* Clarke, D.L. (ed.), 1–60. London: Methuen.

Clarke, D.L. (ed.) 1972b. *Models in archaeology.* London: Methuen.

Clarke, D.L. 1972c. A provisional model of an iron age society and its settlement system. In *Models in archaeology,* Clarke, D.L. (ed.), 801–869. London: Methuen.

Clarke, D.L. (ed.) 1977. *Spatial Archaeology.* London: Academic Press.

Cowen, D.J. 1988. GIS versus CAD versus DBMS: what are the differences? *Photogrammetric Engineering and Remote Sensing* **54(11),** 1551–1555.

Creamer, W. 1985. The upper Klethia Valley: computer generated maps of site location. Paper presented at the 50th annual meeting of the Society for American Archaeology, Denver.

Cribb, R. 1986. Introducing the concept of a dedicated spatial analysis package for archaeology. *Archaeological Computing Newsletter* **9,** 5–12.

Cribb, R. 1987. Some spatial analysis applications programs. *Archaeological Computing Newsletter* **13,** 6–12.

Cribb, R. & M.Minnegal 1989. Spatial analysis on a dugong consumption site at Princess Charlotte Bay, North Queensland. *Archaeology in Oceania* **24(1),** 1–12.

Djindjian, F. 1988. Improvements in intra-site spatial analysis techniques. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 95–106. Oxford: British Archaeological Reports.

DoE 1987. *Handling Geographic Information, Report to the Secretary of State for the Environment of the Committee of Enquiry into the Handling of Geographic Information, chaired by Lord Chorley.* London: Her Majesty's Stationery Office.

Donoghue, D. 1989. Remote sensing and wetland archaeology. In *The Archaeology of Rural Wetlands in England,* Coles, J.M. & B.Coles (eds), WARP Occasional Paper 2, 42–45. Exeter: Department of History and Archaeology, University of Exeter.

Donoghue, D. & I.Shennan 1988. The application of multispectral remote sensing techniques to wetland archaeology. In *The Exploitation of Wetlands,* Murphy, P. & C.French (eds), 47–59. Oxford: British Archaeological Reports.

Ferguson, T.A. 1985. Use of Geographic Information Systems to recognize patterns of prehistoric cultural adaption. Paper presented at the 50th Annual Meeting of the Society for American Archaeology, Denver.

Ferguson, T.A. 1986. The professional Map Analysis Package (pMAP) and its archaeological application. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans,* New Orleans: Society for American Archaeology.

Fox, C. 1932. *The Personality of Britain.* Cardiff: National Museum of Wales.

Gill, S.J. & D.Howes 1985. A Geographical Information System approach to the use of surface samples in intra-site distributional analysis. Paper presented at UISPP Commission IV Symposium on Data Management and Mathematical Methods in Archaeology, Denver.

Grant, E. (ed.) 1986. *Central Places, Archaeology and History.* University of Sheffield: Department of Prehistory and Archaeology.

Harris, T.M. 1985. GIS design for archaeological site information retrieval and predictive modelling. In *Professional Archaeology in Sussex: the next five years.* Seaford, Sussex: Association of Field Archaeologists.

Harris, T.M. 1986. Geographic information system design for archaeological site information retrieval. In *Computer Applications in Archaeology 1986,* 148–161. Birmingham: University of Birmingham.

Harris, T.M. 1988. Digital terrain modelling and three-dimensional surface graphics for landscape and site analysis in archaeology and regional planning. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 161–172. Oxford: British Archaeological Reports.

Harris, T.M. & C.Dudley 1984. A computer-based archaeological site information retrieval system for the Brighton area: a preliminary report. *Sussex Archaeological Society Newsletter* **42,** 366.

Harris, T.M. & G.R.Lock 1991. Integrating spatial information in computerized sites and monuments records: meeting archaeological requirements in the 1990s. In *Computer applications and quantitative methods in archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 165–173. Oxford: British Archaeological Reports.

Hietala, H. (ed.) 1984. *Intrasite spatial analysis in archaeology.* Cambridge: Cambridge University Press.

Hodder, I. 1971. The use of nearest neighbour analysis. *Cornish Archaeology,* **10,** 35–36.

Hodder, I. 1977. Spatial studies in archaeology. *Progress in Human Geography,* 33–64.

Hodder, I. (ed.) 1978. *The Spatial Organisation of Culture.* London: Duckworth.

Hodder, I. & C.Orton 1976. *Spatial Analysis in Archaeology.* Cambridge: Cambridge University Press.

Kintigh, K.W. & A.J.Ammerman 1982. Heuristic approaches to spatial analysis in archaeology. *American Antiquity* **47,** 31–63.

Kvamme, K.L. 1985a. Fundamentals and potential of Geographic Information System techniques for archaeological spatial research. Paper presented at the 50th Annual Meeting of the Society for American Archaeology, Denver.

Kvamme, K.L. 1985b. Geographic Information Systems techniques for regional archaeological research. Paper presented at UISPP Commission IV Symposium on Data Management and Mathematical Methods in archaeology, Denver.

Kvamme, K.L. 1986a. ARGIS: an archaeological Geographic Information System. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans.* New Orleans: Society for American Archaeology.

Kvamme, K.L. 1986b. An overview of Geographic Information Systems for archaeological research and data management. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans.* New Orleans: Society for American Archaeology.

Kvamme, K.L. 1989. Geographic Information Systems in regional research and data management. In *Archaeological Method and Theory,* Schiffer, M.B. (ed.), **1,** 139–203. Tucson: University of Arizona Press.

Kvamme, K.L. & T.A.Kohler 1988. Geographic Information Systems: technical aids for data collection, analysis, and displays. In *Quantifying the Present and Predicting the Past: theory, method and application of archaeological predictive modeling,* Judge, J.W. & L.Sebastian (eds), 493–547. Washington, DC: Government Printing Office.

Limp, W.F. & J.A.Farley 1986. Utilizing computerized mapping and Geographic Information Systems for exploratory data analysis of geobased archaeological data. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans.* New Orleans: Society for American Archaeology.

Lindenberg, R.F. & P.F.Fisher 1988. Towards recognition of the functional distinctions among remote sensing, geographic information systems and cartography. In *1988 ACSM-ASPRS Fall Convention,* ASPRS Technical Papers, 151–159. Falls Church, Virginia: American Society for Photogrammetry and Remote Sensing.

Lockyear, K. & S.Rahtz (eds) 1991. *Computer Applications and Quantitative Methods in Archaeology 1990.* Oxford: British Archaeological Reports.

Madsen, T. 1988. Prospects for the use of the computer in Danish archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S. (ed.), 417–422. Oxford: British Archaeological Reports.

Mayes, M.H. 1986. OS small scales digital development—past, present and future. In *Research Based on Ordnance Survey Small Scales Digital Data,* Visvalingam, M. (ed.), 2–15. London: Cartographic Information Systems Research Group, Institute of British Geographers.

McHarg, I.L. 1969. *Design with Nature.* New York: Natural History Press.

Miller, L. 1986. MIPS: A microcomputer Map and Image Processing System developed for easy use by individual natural scientists. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans.* New Orleans: Society for American Archaeology.

Nagy, G. & S.Wagle 1979. Geographic data processing. *Computing Surveys* **11(2),** 139–181.

Oliver, S.G. & E.K.Schroeder 1986. Archaeological applications of ARC/INFO: a Geographic Information System. In *Workshop Proceedings of Microcomputers in Archaeology, New Orleans.* New Orleans: Society for American Archaeology.

Rahtz, S.P.Q. (ed.) 1988. *Computer and Quantitative Methods in Archaeology 1988.* Oxford: British Archaeological Reports.

Rhind, D. 1977. Computer-aided cartography. *Transactions of the Institute of British Geographers* **2,** 71–97.

Rhind, D. 1981. Geographical information systems in Britain. In *Quantitative Geography,* Wrigley, N. & R.J.Bennett (eds). London: Routledge.

Rhind, D.W. & H.M.Mounsey 1989. Research policy and review 29: the Chorley Committee and ‘Handling Geographic Information’. *Environment and Planning A* **21,** 571–585.

Rodgers, P. 1988. Multi-response permutation procedures. In *Computer and quantitative methods in archaeology 1987,* Ruggles, C.L.N. & S.Rahtz (eds), 45–54. Oxford: British Archaeological Reports.

Star, J. & J.Estes 1990. *Geographic Information Systems: an introduction.* New Jersey: Prentice Hall.

Tomlin, C.D. 1990. *Geographic Information Systems and Cartographic Modeling.* New Jersey: Prentice Hall.

Tomlinson, R.F. 1972. *Geographical Data Handling.* Ottawa: International Geographical Union Commission on Geographical Data Sensing and Processing, Ottawa.

Tomlinson, R.F. 1987. Current and potential uses of geographical information systems: the North American experience. *International Journal of Geographical Information Systems* **1,** 203– 218.

Tukey, J.W. 1977. *Exploratory Data Analysis.* New York: Addison-Wesley.

Wansleeben, M. 1988. Geographic Information Systems in archaeological research. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 435–451. Oxford: British Archaeological Reports.

Whallon, R. 1984. Unconstrained clustering for the analysis of spatial distributions in archaeology. In *Intrasite spatial analysis in archaeology,* Hietala, H. (ed.), 242–277. Cambridge: Cambridge University Press.

Zubrow, E. 1987. The application of computer-aided GIS to archaeological problems. In *Proceedings of the First Latin American conference on Computers in Geography, San Jose, Costa Rica,* 647–676. Costa Rica: Editorial Universidad Estatal a Distancia.

10

The display and analysis of ridge-and-furrow from topographically surveyed data

MIKE FLETCHER & DICK SPICER

Archaeological surveys

Topographical surveys of earthworks such as hillforts, barrows or ditched enclosures are often conducted by archaeologists with the intention of producing plans, either for publication or for archive purposes. Usually such plans will use the process of hachuring (which uses tadpole-like lines whose thickness varies according to the height represented) to depict local variations in terrain; with skill, hachures can result in elegant and valuable plans, though when they are crudely drawn they are rather ugly. Sometimes plans will incorporate contours where overall slope is significant in some way. The hachure method is purely subjective, the result of interpretation (albeit at an early stage); contours are rather less so, though may be the result of judicious 'improvement'. However, three-dimensional information of any real value to *direct* interpretation is conspicuous by its absence from most archaeological publications, even though it may be obvious from the text descriptions that shape is vital to the interpretation of the site. Nevertheless, there are signs of an awareness of the need for the incorporation of better graphical techniques with the revival of the previously rather frowned upon method of hill-shading (simulating shadow from oblique lighting) in recent earthwork publications (for example Royal Commission on Ancient and Historic Monuments—Wales 1988).

Occasionally, surveys of earthworks have been published as 'wire' or 'net' diagrams (for example Barker 1986, p. 62; see also Lock 1980) when the data from the survey is of a sufficient density and regularity to permit this. The effort put into a survey, and hence its ultimate value, is usually determined by the complexity of the site as perceived by the archaeologist at the time of the survey. The authors have outlined a method by which trial surveys at different densities might be examined for their value in representing the whole area (Fletcher & Spicer 1988c). Unfortunately, perceptions can change during investigation, and retrospective surveys to fill in lost detail (due to incorrect choice of method) will probably not be possible if the site is excavated.

It is an unfortunate general fact in archaeological fieldwork that surveys are taken only of features which are identified in the field: that decisions are taken on the basis of what is *already known,* not what *might be discovered* from a survey. So medieval earthworks are booked with readings taken from the 'significant' points (which are subjectively chosen) of each identifiable feature. Banks may be represented by the highest points taken along their length, and by their footing with the surrounding terrain. The data obtained from such work is usually quite adequate for the sort of plan intended for publication, but seldom tells the fieldworker anything more than he or she knew before planning. The possibility is overlooked that there may be features present which are invisible to the fieldworker.

> The average site that the medieval field archaeologist will need to survey is not large… This is quite different from a 20-hectare hill fort, with all the problems of long distances and *dead ground* that such a site might involve. (Taylor 1974, p. 38, our italics)

The use of Electronic Distance Measuring (EDM) equipment has not significantly improved matters, other than to reduce by a large factor the amount of time spent actually performing a survey, and (as a relatively unimportant side-effect) enhancing its absolute precision. Such is the pressure under which archaeological fieldworkers do their job that the luxury of a more detailed EDM survey performed within the same time as a manual one is foregone in favour of a rapid but minimal one. The extra time involved in making a thorough and systematic survey of a large area is rewarding when the data is displayed using the right kind of advanced graphical techniques because of the likelihood that more information can be discovered as a result.

Computer graphics and archaeology

Powerful computers can now be used to generate stunningly beautiful models of structures and buildings reconstructed from archaeological data. The aim might simply be to illustrate; but often there may be some more purposeful intention— to investigate or to stimulate thoughts towards an archaeological interpretation. As a tool for scientific analysis, the visualizing process resulting from solid modelling can sometimes reveal relationships within an archaeological ‘reconstruction’ more clearly than other current methods of display. However, the programs which archaeologists often adopt are not designed specifically for this purpose. They are usually intended to work equally well whether the output is a new car design, an atomic structure, or an architect’s modern building. Moreover, when archaeological data is the source, solid modelling will usually entail a high proportion of subjective judgment, since so many controlling parameters are not completely known. Colour, texture, and to a certain extent, dimensions, are frequently a matter of choice made by the programmer/archaeologist, based upon assumptions or knowledge obtained from other sources. Solid modelling (see Wood & Chapman, (with contributions by Delooze & Trueman) 1992; Reilly 1992) by its very nature encourages the copying of the shape segments which make up the structure of a solid-modelled building; columns, for example, are usually constructed once, and repeated over and over again wherever needed. Any offending visual repetition which this might produce can often be removed by discrete use of texture mapping and lighting variations. The final result is a sort of architect’s perfect vision, which in the case of a medieval building, such as an abbey, can still be extremely valuable to the archaeologist. As with all reconstructions, however, the observer might not be fully aware of how much of a particular picture was conjectural, since the whole image is ‘perfect’. This dilemma of ‘too much of a good thing’ is being considered now by some, and parts of the model which are questionable are being shaded in a different way, rather in the manner of the notorious ‘pink cement’ which was once used in Britain by the state monument guardians to mark repairs on real sites.

Prehistorians are seldom able to unearth data of sufficient quality or detail to permit a reconstruction suitable for these type of solid models. Indeed prehistoric structures of any kind are scarce, since the materials used were either not substantial (wood rarely survives), or uniform (only one of the many hundred British stone circles used artificially shaped stones), and not therefore of an easily modelled nature. A recent solid model shown on national television of the newly-discovered timber enclosures south of Avebury was effective in demonstrating their scale when superimposed on a view of the site as it is today, even though each of the posts was just the same shape as its neighbour. It is interesting to note that when prehistoric stone structures such as Stonehenge are solid modelled (the front cover of *Computer Bulletin*, September 1986) the results are of a very stylised quality, and convey nothing to a prehistorian, who would be concerned with shape and style of the individual stones.

Modelling the shape of earthworks from actual survey, on the other hand, is valuable because the earthworks are themselves the structure—albeit in a denuded state—intended by the builders. Prehistoric monuments (Colour Figure 10.1) and features such as henges, barrows, hillfort banks and ditches, and dark age and medieval earthworks are all surfaces which, although no doubt originally supporting superstructures of an ephemeral material, can still provide valuable information as to their purpose. Aerial photographs have for decades been a productive means of identification and discovery of earthworks which are too denuded to be seen from the ground, making use of low angle illumination from the sun, or differential snow cover, or crop colouration to reveal and delineate subtle shapes. Plans taken from interpretation of such aerial photographs can be rectified geometrically (eg. Haigh 1989) to tie features into maps, frequently supplying much valuable information from just one picture.

It was implied above that there is a difference between reconstruction—that is, the production of a complete and *finished* image from minimal data—and the *innocent representation* of data gathered by various survey methods. Since many of these latter methods set out to explore surfaces, the subject of this chapter will be the representation of surfaces rather than anything else. The object of representation of a surface is the communication of information in the most efficient manner possible. The viewer is asked to visualize a *real* scene. The fact that the scene may not actually be physically viewable is irrelevant—the ‘site’ may have been destroyed, it may be on the other side of the world, or completely covered in trees; it may even be an ephemeral archaeological layer, or a ‘pseudo-surface’ derived from electronic geophysical survey. What, however, is common to all data of this kind is that it can be *visualized* as a real, tangible surface. A *papier mâché* model could be made of it and held up for everyone to see, though in the case of topographical survey data it would require a good deal of painting and careful positioning of a light source even to begin to look like the real thing.

Elementary though this may seem, we need to remind ourselves just what it is we are viewing. A radio programme a few years ago described a computer-generated model of archaeological data as looking like something from ‘Star Wars’. Was this a compliment or ought it be taken as a criticism? Could the lay human observer gain information from the coloured image on the screen or was it something alien to be interpreted only by eyes which are accustomed to weird shapes and colours? It is our objective to produce on the screen something which is as realistic as possible, without cheating, fudging or tampering with the data in any way during the display process. The criterion is simply to show the surface of the ground in as naturalistic a way as possible, so as to utilize fully the human abilities of observation and perception, and to couple it with the experience and intuition of an archaeological training. The surface modelling has been written from the outset specifically for the display of archaeological information—not for molecular structures, television programme titles or anything else. Its limitations are apparent; but it is growing and developing as better hardware becomes cheaper, and will accommodate newer methods as time permits. The project has been exploratory, with the intention of investigating just how much information can be presented visually.

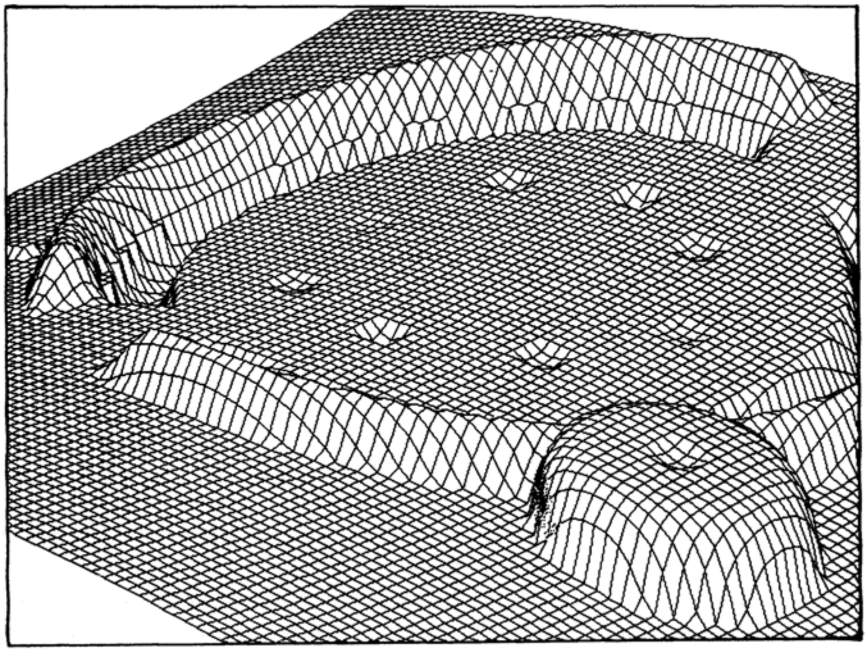


Figure 10.2 Clonehenge artificial site—wire diagram (200×200 points).

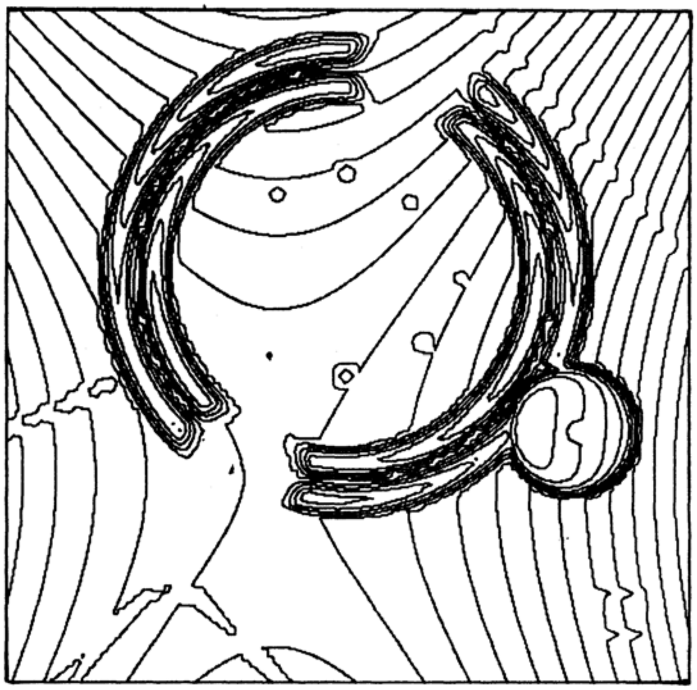


Figure 10.3 Clonehenge contour plan (as Figure 10.2).

What has become very apparent during these investigations is that *interaction* by the archaeologist with the machine is a vital element to the investigation process. The publication of one or two pictures is quite inadequate as a description of the data, since so many new aspects can be discovered by varying the viewing parameters. Those who are devoted to the idea of electronic interactive publishing should perhaps take note: it is the ability to view from all angles and distances, under a variety of lighting conditions and with as many colour controls as possible which brings about rewards. Reilly has referred enthusiastically to the value of simulating aerial reconnaissance (Reilly 1988a, ch. 8), where *pseudo-lit* wire frame models can be moved in real time by manipulation of tracker devices, but this was performed with the backing of vast hardware resources

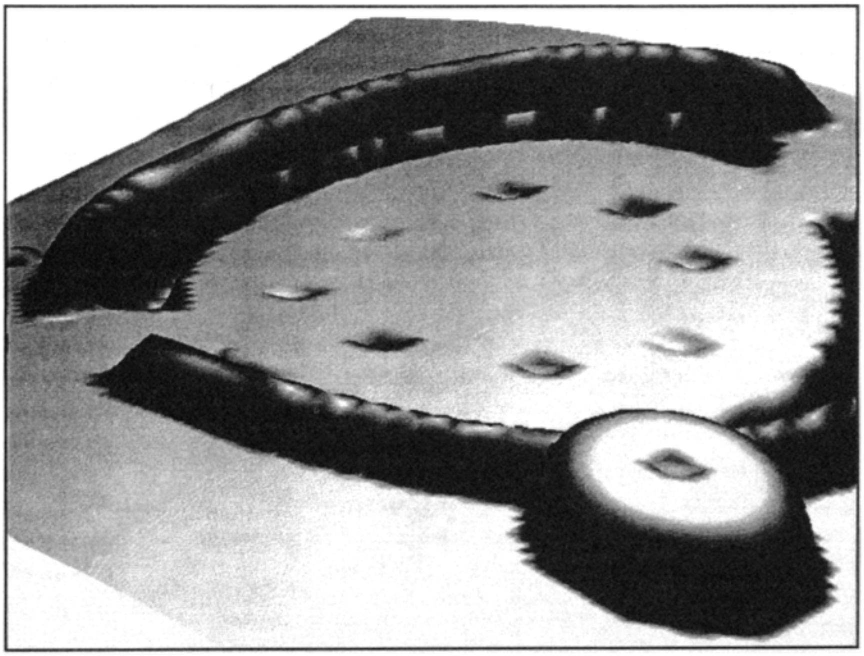


Figure 10.4 Clonehenge—lit surface model (as Figures 10.2 & 10.3).

beyond the reach of most archaeologists. Our fully lit surfaces take rather longer (but typically under half a minute on a PC) to generate, and the rewards are just as great as if they were instant. The images can, of course, be stored and viewed rapidly in a 'slide-show' sequence, though the viewer then forgoes the luxury of controlling precisely what he or she sees. That the program has proved as addictive as any good computer game to all archaeologists who have used it is some indication of the value of interaction.

Just as aerial photography benefits from a variety of camera angles and (to a degree limited by the weather) choice of lighting quality, so the viewing and lighting conditions of the computer model should be flexible and interactively controllable by the archaeologist. For this to be possible, the software must be reasonably fast and easy to use, and the hardware of a modest enough specification to be available to the majority of budget-conscious field units. It was therefore necessary to design the graphical programs to work on current PC configurations, since these are in widespread use, without limiting too much the results by any hardware inadequacies. VGA screens are now improving rapidly both in resolution and colour control, and standards are likely to improve along similar and (it is to be hoped) compatible directions for the foreseeable future. Currently the mandatory minimum is the 'standard' VGA screen of 640×480 pixels, using normally 16 colours. 256 colours are used for four-dimensional surfaces (see below). A fast machine (eg. a 386 with maths co-processor) is desirable but by no means essential.

Methods of surface modelling

The object of this kind of graphical application is to provide a description of the archaeological data in a way which is optimised to communicate the maximum of information, with as little spurious addition as possible. Although we believe that the naturalistic approach is the ultimate, we do not discard the conventional contour plans or wire diagrams when developing new methods, which still have their place in the archaeological report, notably because they are easy to produce and lend themselves to good quality cheap reproduction. Just as a variation in angle and lighting can trigger perceptual clues, so can the use of diagrammatic forms of display. Contours can sometimes show subtle shapes; they can even be drawn in three dimensions, though they seldom provide very much information unless used in combination with other methods (Lock 1980, p. 24; Spicer 1985, Fig. 6). With this in mind, any modelling system should be capable of providing the whole gamut of display methods, each of which might have a particular virtue or an ability to help identify one or more characteristics. No single graphical technique so far produced has been able solely to provide a 'total answer'. The range includes familiar traditional methods such as the contour, the wire diagram and extends through to three- and four-dimensional surface models (see Figs 10.2–10.4).

The wire diagram has the virtues of speed and ease of reproduction as a monochrome line print. As a minimum requirement, however, it should take into account perspective, but not necessarily hidden lines. With a gridded wire surface, the need for an integral scale is rare, the lines themselves providing it, and it may detract from the perception process.

Similarly the *plinth* (the frame or base on which the 'surface' sits) is seldom essential, since cliff-edges do not form part of the original data. The ability of the human perceptive process to 'fill in' the panels within a grid is often underestimated: indeed mere dots are all that is necessary to stimulate the brain into seeing a surface (Spicer 1985, Fig. 5). Wire diagrams can suffer badly from perceptual problems, however, notably the Necker Cube phenomenon and eye-fatigue effects due to the repetitive nature of the diagram (Spicer 1985, p. 14). Colour banding can improve wire frames further and so help in removing ambiguities, whilst a 'pseudolit wire surface', in the manner proposed by Reilly (1988a, ch. 8), can be an easy enhancement to an already simple method. Wire frames, however, are only capable of showing *overall* terrain shape, and the small surface irregularities which might reveal features to an experienced archaeological eye are usually imperceptible, even though they might be present in the data. Contours can still be important in emphasising or delineating features where relationships could be ambiguous, or need to have their position indicated precisely. The main disadvantage with contours is that they seldom can reveal both overall terrain and local anomalies at the same time. Contours of an archaeological earthwork, such as a ditched enclosure, situated on the gentlest slope will be masked by the contours of the overall terrain. Moreover, although we now accept them as a conventional way of depicting topology, the actual image of squiggly lines has no familiar analogue: it is so unnatural, never encountered in reality, requiring an active intellectual process between perception and interpretation.

When displaying three-dimensional vector information, whether in the form of contours or wire-frames, the choice of whether or not to hide lines which would be obscured by a surface is not an easy one, since information will certainly be lost by hiding lines, and interpretation might be confused by not hiding them. Stereoscopy can be called upon to try to help remove ambiguities of this kind, though the impracticalities of viewing and reproduction can outweigh the advantages (Spicer 1985, p. 16). Adding solidity (or rather, opacity) to the surface makes the view more credible perceptually, and even simple panels, coloured according to height, give impact to a diagram. This only works well when the gridded model is itself of high definition, that is, the grid size is small compared with the whole area. The picture at once becomes more lively, though the shading provided by hard copy devices (in the form of regular dither patterns) is often a very poor reproduction of a coloured screen display. Lighting the panels can reveal details and emphasize small features which other methods might miss. Their value in highlighting small features has been discussed elsewhere (Spicer 1988). Though not by any means a realistic model, lit panels are still very worthwhile.

Smooth lit surfaces can provide very acceptable degrees of realism, often simulating an aerial photograph. The lighting is from an artificial sun, infinitely distant and freely movable, allowing the observation of particular times of day to be produced. Multiple light sources, of the kind developed by many solid and surface modellers, is not desirable in this case, since it would produce an unnatural effect. The surface model can also provide the base for the addition of some of the other methods listed above—lines in the form of grids or contours can be superimposed, or the colour can be made to alter according to some other kind of data—such as soil type, occupation phase, density of pottery sherds and so on, as a fourth dimension. It might be the outcome of another type of survey—a geophysical one (Arnold, Huggett, Reilly & Springham 1989, Fig. 16.2)—or even be the consequence of some sort of manipulation or filtering of the original data. However, these should not be allowed to detract from the perception of the physical shape of the image. In all cases we feel it is desirable to maintain realism by not sacrificing lighting of the primary model. This usually means an increase in the number of screen colours available, though it is possible to use dithering techniques for intensity changes instead, though at the sacrifice of screen resolution. The importance of this point perhaps needs emphasising, since a number of CAD (Computer Aided Design) and GIS (Graphical Information Systems) packages now provide four-dimensional facilities (for example Richards 1991) but do so often at the expense of lighting so that the result is perceptually nothing more than a coloured wire model, requiring grid wires to supply the three-dimensional information. An example of a lit four-dimensional model is given in Colour Figure 10.5.

Viewing a data set with the wide variety of methods described above will quickly show the virtues of a realistic approach to display, since it is from a lit surface that the best impression is obtained, and the natural human perceptive processes can work most efficiently. It requires little by way of interpretation (in the non-archaeological sense), making perceptual sense even when 'impossible' sun positions (such as from the north in northern latitudes) are invoked. Choice of colour palette is important in this case: a simple gradation of monochrome loses over a variation of **both** intensity and hue, as shown in Colour Figure 10.6, where the hue changes from dark blue, through shades of green and yellow, and finally to white. As with all graphical techniques, experimentation will frequently reveal features within the model which are of archaeological interest. Use of palette look-up table changes can trigger perception of particular features which may not be immediately obvious. Reversing a lighting palette can produce the appearance of an overcast day—a sort of 'non-sun'— when the model is lit from all directions except one, but this needs a uniform-toned palette in order to avoid problems with the highlights artificially produced with tonal changes. The palettes can be ordered in some way—and may be variations on a natural sequence—or can be random in order, producing a 'false colour' effect, where a particular colour scheme may have a serendipitous result. Though not to be recommended as an end in itself, the use of false colours can often be rewarding when used in combination with a natural palette, which should always be used for any final interpretation.

Movement can often stimulate perception, and this can sometimes be simulated by rapidly switching palettes, alternating highlights and shadows. Ridge and furrow is well displayed in this way. Since images can be stored and later retrieved rapidly

— typically about one a second—movement can be produced from a sequence, which may include changes of all parameters, including sun positions.

In addition to providing numerous methods of viewing archaeological surface data, a certain amount of careful filtering of the data itself has proved most revealing. Use of standard, well-known image-processing techniques, such as local-averaging window convolutions can expand possibilities of presentation methods. It must be stressed, however, that these processes work best in small doses at a time: that is, each stage in any data manipulation should be monitored using the same high-level graphical output. Combining a series of processes, or using a large window rather than multiple passes of a small one, might well deprive the researcher of valuable 'in-between' pictures. An earthwork site might slowly be 'eroded' by a sequence of low-pass filters. To an archaeologist, the idea of being able to produce an approximation of an 'old ground surface'—that is, the terrain before a structure was built—is very attractive. This can easily be done using a local average filter, repeating the process until features of a certain subjectively chosen size are gone. Having produced this surface, it is an easy matter to derive the residuals (by subtraction), so that the archaeologist can see an entire site 'flattened'. By using one of a number of gradient filters, features which orientate in a particular direction can be identified, as illustrated below. All the above techniques have variously been employed in connection with the subject of the second part of this chapter, an examination of an area which has archaeological features of a number of different origins.

Surveying strategy

In addition to researching graphical methods of representing archaeological earthwork data the authors have found it necessary to examine the surveying methods currently in use. It is our contention that a thorough and systematic ground survey can frequently reveal much more than might be seen on the ground, even after the close familiarity from extended walking back and fro which such a survey must entail. A computer simulation of an archaeological site, nicknamed *Clonehenge*, was developed and made freely available (Fletcher & Spicer 1988b) in order that the prospective surveyor might try out different strategies, and that different surface-fitting, contouring, and display methods might be systematically compared (Fletcher & Spicer 1988a). Figures 10.2 and 10.3 show the 'site' surveyed on a regular grid of 200 points square, displayed as a wire diagram and a contour plan. Figure 10.4 is a lit surface model of the same data.

Ridge-and-furrow

Despite its familiarity to archaeologists due to its occurrence over most regions of Britain, the regular pattern of strips known generally as ridge-and-furrow has not received a great deal of analysis. It is widely assumed to be due to cultivation during the middle ages; much of the current thinking is covered by Taylor (1975) and Reynolds (1980). When found closely associated with prehistoric sites, it has often been ignored, or at best dismissed without further comment as medieval, in a manner which disregards any possibility of an earlier origin. Neolithic and bronze age plough marks (though not cultivation strips) have been acknowledged and accepted in other contexts, such as under prehistoric burial mounds (eg. Evans 1972).

Nevertheless, there are signs of recent renewed interest. Bronze age field systems, visible by their boundaries, are becoming more fully understood following major research projects of the kind pursued in Dartmoor (Fleming 1988), and now in the Peak District (Barnatt, *pers. comm.*). Such research now shows the extent of the systems to have been greatly underestimated. A recent paper (Topping 1989) describes fieldwork in northern England which shows plans of areas of pre-roman cultivation in areas where pastoral activity was thought to dominate. The 'cord-rig' pattern described is considerably smaller in pitch than the so-called medieval ridge-and-furrow, but its identification both from aerial photographs and on the ground suggests that it might be quite widespread, and that many other highland zone areas should be investigated.

Whilst we are not in any way offering at this stage *any evidence* of prehistoric ridge-and-furrow, it is nevertheless clear that at certain sites, detailed analysis of the juxtaposition of field boundaries, cairns and cultivation ridges could reveal a history of cultivation over many discrete periods, and might explain certain anomalies of temporal phasing.

Stapeley Hill

A bronze age site surrounded by quite clear ridges (Colour Figure 10.1) and never before investigated, was suggested by Mike Watson, archaeologist for Shropshire County Council, to be a candidate for examination and a test-bed for some of the techniques which we were developing. It is from this site that the results presented in the remainder of this chapter have been taken. It is situated near the border between England and Wales, not far from the town of Montgomery, in an area well-known for prehistoric monuments. The site in question, a round burial mound known as a ring cairn is marked on the Ordnance Survey 1:2500 map at National Grid Reference SO 3127 9904 on a saddle-shaped ridge, known as Stapeley Hill, which runs roughly NNE-SSW, an azimuth of about 30° to grid north. The hill is quite straight, level, and steep sided; the gradient is as much as 7° from horizontal on either side. The cairn is placed precisely on the middle of the saddle, at a height of approximately 398m OD (ie. above sea-level). The highest point of the hill (at 403m) is about 150m to the NNE and another slightly higher point 85m to the SSW of the cairn.

The sites and monuments record (SMR) states that the cairn itself has been robbed of stone, and gives the height at its centre as 0.9m and the diameter 15m. More specifically it records the internal mound as being 7.0m diameter and 0.3m high; the width of the bank is given as between 2.5m and 5m, with heights of up to 0.6m external and 0.4m internal.

The whole area of Stapeley and its surroundings has many field systems which cross the saddle hill roughly orthogonally, and it is these systems which continue in the lower regions southwards around the well-known stone circle called Mitchell's Fold. A fine aerial photograph (reference 87-MB-73) taken by Chris Musson in 1987 shows the whole of Stapeley Hill from the east; figure 10.7 is an enlargement of part of this photograph. The ring cairn is very clear, and the many and varied ridges can clearly be seen crossing the hill.

A detailed survey was made using a dumpy level, with a pocket computer (a Psion Organiser II) used for data entry and on-site validation, on a carefully laid-out grid of half-metre intervals, the angle of which was chosen in accordance with the results of the simulation detailed below. Ridges in the area repeat at approximate intervals of 2.8m, though with considerable variability. Sometimes, a particular ridge may be 'lost' (or omitted), resulting in a dubious furrow of double or even triple width. The survey interval was chosen as a good compromise between survey efficiency on the one hand, and lack of detail on the other.

It was felt important to be thorough in producing the survey, since we were producing test data for methods yet untried. Crude tests had shown that a *single* traverse performed at right angles to the ridges (so as to cut them) could not be expected to sort out the many sources of noise (due to vegetation, sheep-tracks, minor survey errors, and so on) which was present in the data. Moreover, the direction of the ridges was slightly different on either side of the hill, and assessing on the ground the precise angle of each ridge is extremely difficult, since a subjective judgment has to be made in determining the crest or trough of each ridge or furrow. A gridded survey taken over a substantial distance would provide a series of traverses across each ridge. This would provide the data for mathematically establishing the angle of the ridges.

The exhaustive tests performed on simulated ridge-and-furrow data which followed, described by Fletcher & Spicer (1989), revealed that not only the pitch of the survey sampling points, but the angle between the survey grid and the ridges themselves was crucial to avoid the most severe effects of aliasing. Angles very close to 0, 90° and even 45° produced interference which was of a level high enough to alter drastically the result at certain points of an angle-detecting algorithm. The conclusion, based empirically on evidence from crude analysis, was that angles of 20 or 30° were acceptable. Because the Stapeley ridges appeared to alter their orientation slightly over the area examined, no attempt was made to measure one single global direction for the whole site, but rather to estimate the angles at all points of the grid using a small moving window centred around each of the grid points in turn.

Because of its position in the midst of the field systems, the ring cairn was chosen to be central to the survey; around it, and running down the slopes to either side, were ridges sufficiently clear on the ground in a number of places, their relationship to the cairn (that is, whether they respected its presence) being of obvious interest. The value of the choice of Stapeley as a test site was as follows:

- Ridge and furrow was visible at the periphery and beyond the surveyed area.
- There was some uncertainty about the ridges at the interface with the cairn.
- There was a visible change of direction of most ridges at the highest point.
- A denuded field wall, possibly earlier than the ridges, was adjacent, and could be included in the survey.
- The whole site had not been surveyed before, and the new data was of general archaeological value.

The surface model (Fig. 10.8) shows that the ring cairn is right on the crest of the hill, with fairly steep slopes on either side. Each intersection in the wire diagram represents a single height reading, with a horizontal distance of 0.5m between each adjacent reading. The data has been low-pass filtered slightly to remove minor survey noise, and the vertical scale increased by two.

The lit surface model (Fig. 10.9, and also Colour Figure 10.6) shows the ring cairn very clearly, and the shape of the topography in which it is situated. Some ridge-and-furrow can be seen crossing the hilltop, and the field wall can be seen on the right about halfway across the site. Colour Figure 10.1 is a photograph of the real site taken from approximately the same angle as the model in Colour Figure 10.6.

Analysis of linear ridges

Figure 10.10 is a contour map of Stapeley Ring Cairn, and Figure 10.11 is of the filtered residuals. Although the ring cairn is plainly visible, the ridge-and-furrow is not clear. The residual contours do not show the overall topography, which of course falls away from the ring cairn towards bottom right and top left. The close contours at the top of the picture show a denuded

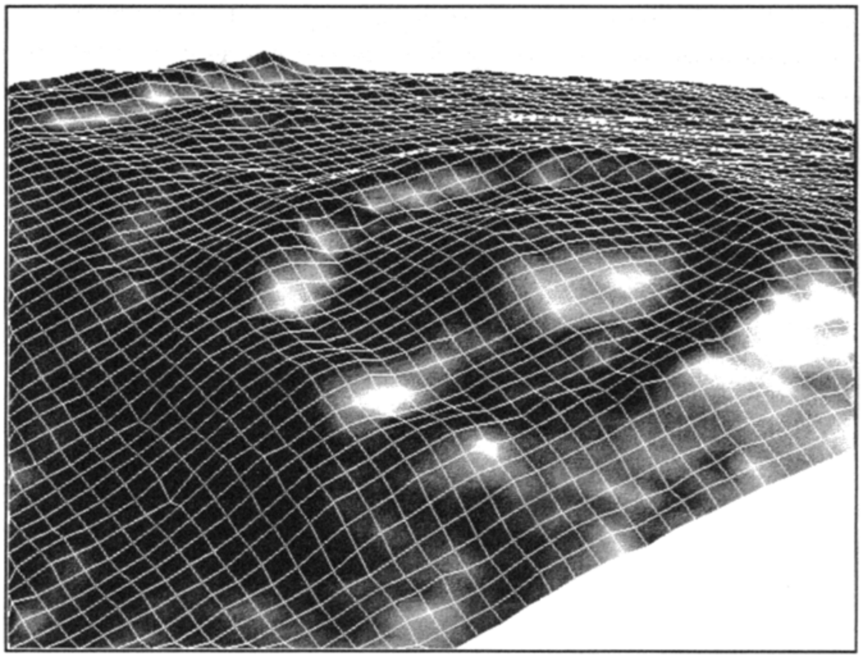


Figure 10.8 Stapeley Ring Cairn—close-up of cairn.

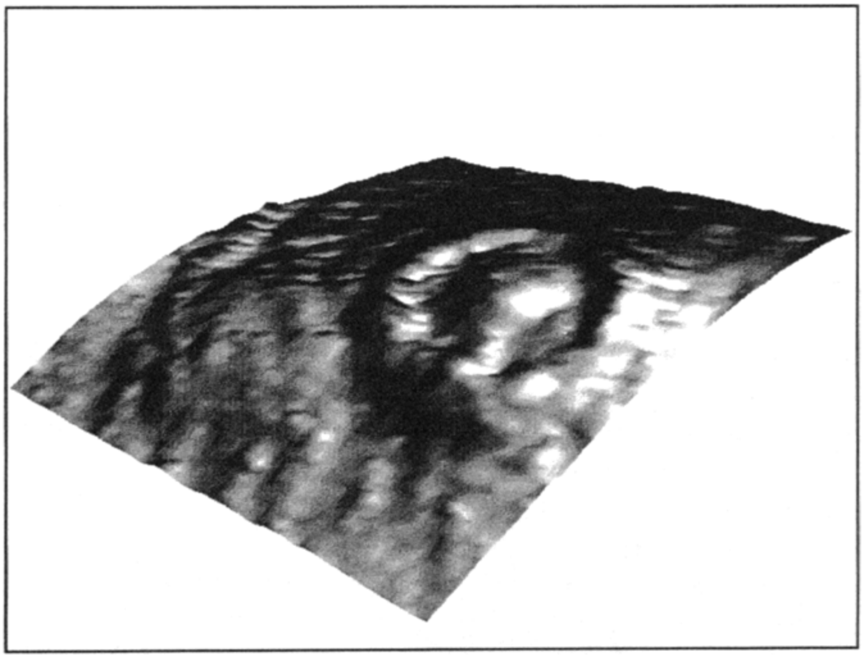


Figure 10.9 Stapeley Ring Cairn—lit surface.

field wall and there is a hint of a feature, which could be another mound or might be an avenue entering the ring cairn from top right. Identifying the ridge-and-furrow is difficult and rather subjective.

One immediately obvious technique in tackling the problem was to try to filter the data so as to enhance any directional trends and suppress other features. A simple method was to apply variations on the Sobel gradient detection filter (Fig. 10.12). The results can be displayed as a fourth dimension on top of the topographical model (Colour Figure 10.5), where the colour extremes of red and blue indicate negative and positive gradients, and the intensity still indicates lighting. Although the results were promising they are accompanied by the spurious effect of shifting the phase of the ridges, and, above all, the method provides no real quantification of confidence of the presence of ridges. Another approach was to estimate the angle of the ridge-and-furrow by weighted regression of the raw data. This immediately revealed another problem: the underlying topography of the site masked any ridge-and-furrow present. So instead of using the raw data, residual data

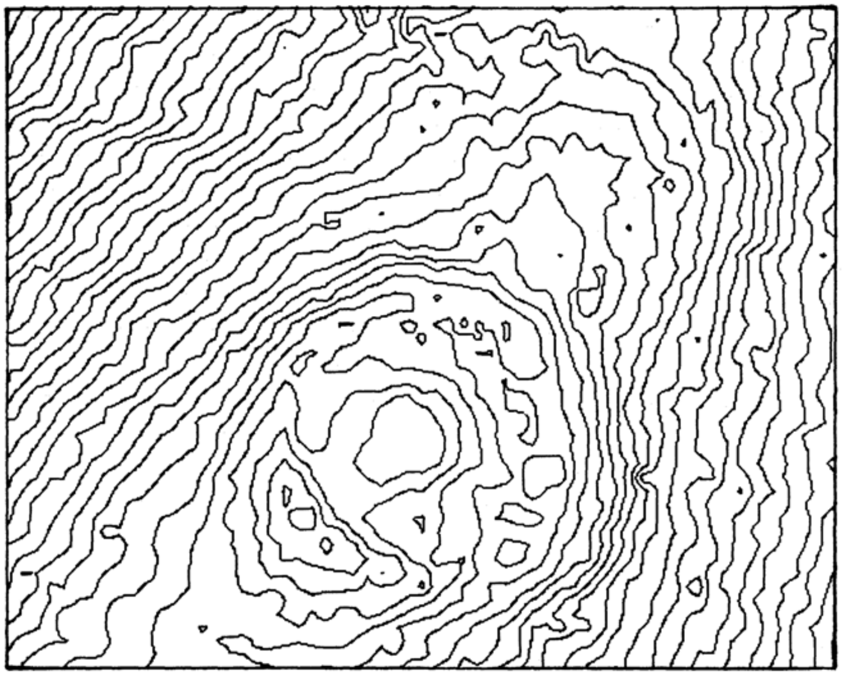


Figure 10.10 Contour plan of Stapeley Ring Cairn.

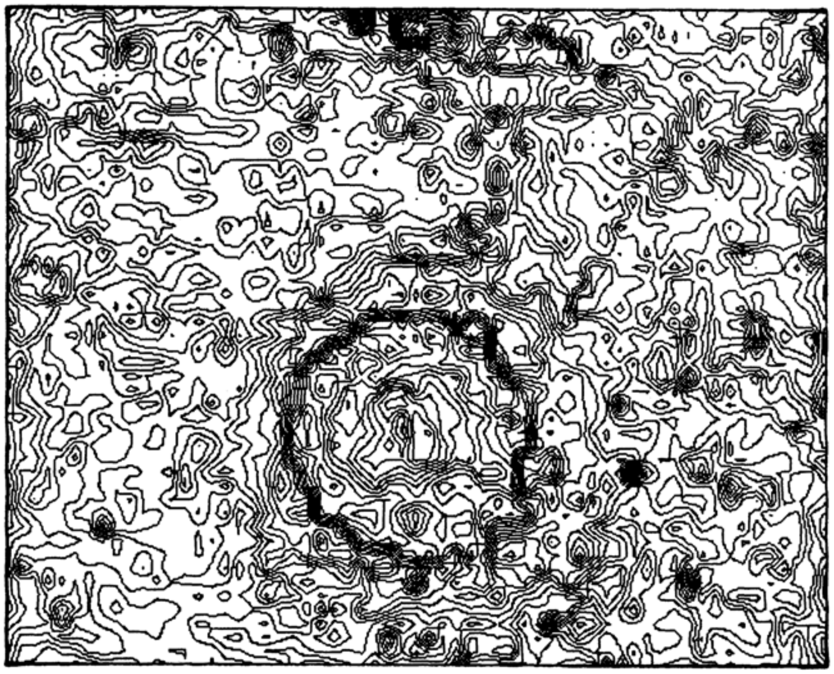


Figure 10.11 Contour plan of residuals after filtering.

was used obtained by differencing the original from a suitable multiply-smoothed version. Although this produced some interesting results, there were still some effects caused by any remaining slope trend in the residuals and so a more refined and versatile method was sought.

A method of measuring orientation suggested by Zucker (1987) is very costly on machine time and is essentially a contouring technique. The Radon transform (Dougherty & Giardina 1987, pp. 217ff) involves integrating the data or image along lines of different gradient θ and different perpendicular distance from the origin p and thus transforming the image $f(x, y)$ into $R(p,q)$. Although after some manipulation of the transformed data this method gives an indication of any favoured value of θ it was not developed further because it did not have any means of distinguishing between different sizes of ridge-and-furrow.

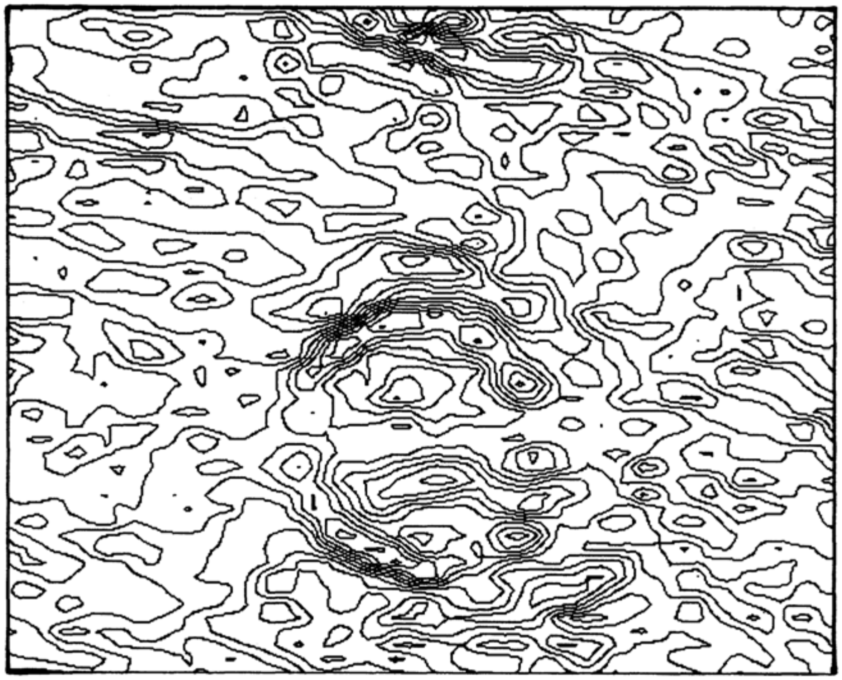


Figure 10.12 Stapeley Ring Cairn after Sobel filter.

The technique presented here involves passing a ten-by-ten grid-points window over the residual data and estimating both the angle and amount of confidence of any ridge-like pattern within the window. Ten as the dimension was estimated as the optimum for data surveyed every 0.5m and with ridges approximately 3m apart. Figures 10.13 and 10.14 show the sample data sets used to demonstrate the method, each being 100 residual values $f(x, y)$ on a 10×10 grid. In both cases the ridge-and-furrow is at 70° from (nominal) north in a clockwise direction, Figure 10.13 having ridges 2.8m apart, and Figure 10.14 having ridges 3.6m apart.

When these data sets are transformed into the frequency domain (u, v) to obtain the power function $P(u, v)$ the results shown in Figures 10.15 and 10.16 are obtained.

These clearly yield sufficient information to estimate angle of the ridge-and-furrow. Although a Fourier transform would produce similar results, a Hartley transform modified to produce just the power function was finally used since it was considerably quicker. The power images produced are optically transformed (that is, each quadrant of the image is exchanged diagonally) and the central value (a very large value corresponding to the constant term) has been put to zero to produce a clearer picture. If the data on the 10×10 grid is denoted by $f(x, y)$, with $0 \le x, y \le 9$ then the Hartley transform of $f(x, y)$ is $H(u, v)$ where

$$H(u,v)=\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} f(x,y)\{\cos[(ux+vy)2\pi/n]+\sin[(ux+vy)2\pi/n]\}/n^2$$

and the power $P(u, v)$ is given by

$$P(u,v)=[H(u,v)^2+H(-u,-v)^2]/2$$

To estimate the slope of a straight line through the origin (shifted to the centre) passing through these two symmetrically placed peaks in the power function, a weighted regression was used to estimate θ so as to minimise

$$\sum d^2 w$$

where

$$w=\sqrt{[P'(u,v)]}$$

$P'(u, v)$ is the swapped version of $P(u, v)$, and d is as shown in Figure 10.17.

By altering the weights W so that only those powers that lie inside a chosen diameter are used, the data can be suitably filtered. At this site we are looking for ridge-and-furrow just under 3m apart and we found the appropriate cut-off was achieved by using those values of $P'(u, v)$ for which $u2+v2<9$.

The value of θ which produces this minimum can be shown to be:

$$\theta=0.5\tan^{-1}[2\sum uvw/(\sum u^2w-\sum v^2w)]$$

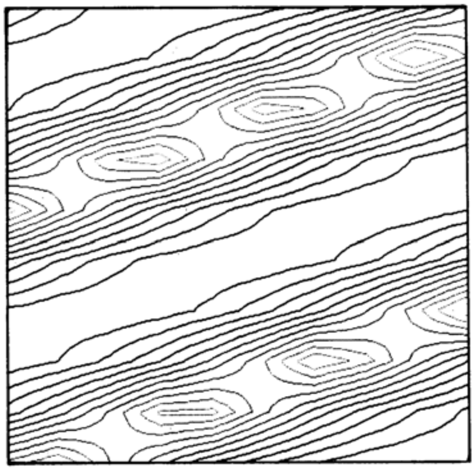


Figure 10.13 Contours of Ridge and Furrow at 70° 2.8m apart.

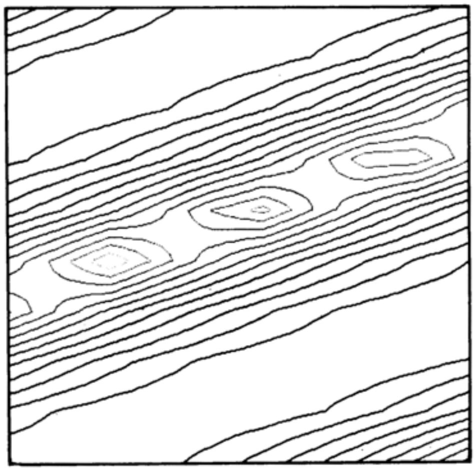


Figure 10.14 Contours of Ridge and Furrow at 70° 3.6m apart.

One measure of confidence can be defined by:

$$C = 1 - \min_{\theta}\left[\sum d^2 w\right] / \max_{\theta}\left[\sum d^2 w\right]$$

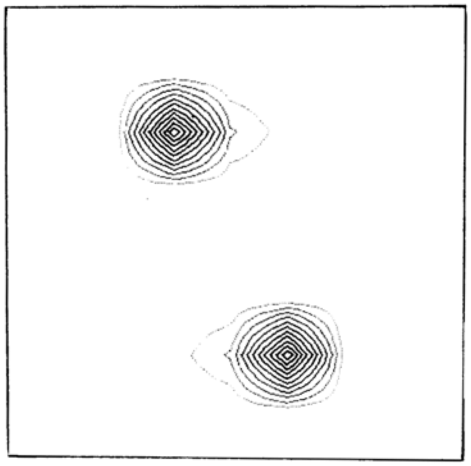


Figure 10.15 Power Function 70° 2.8m apart.

Figure 10.16 Power Function 70° 3.6m apart.

This will yield values close to zero when the data has no ridge-and-furrow but will provide values near to 1 when the ridge-and-furrow is clear. When different sets of test data for ridge-and-furrow at varying angles θ were used it was found that C was not altogether independent of θ, as it ought to have been. A better measure of confidence is obtained by using

$$\text{Confidence} = \sqrt{\frac{C}{(0.7C' + 0.3)}}$$

where C' is the value obtained for C but with no filtering.

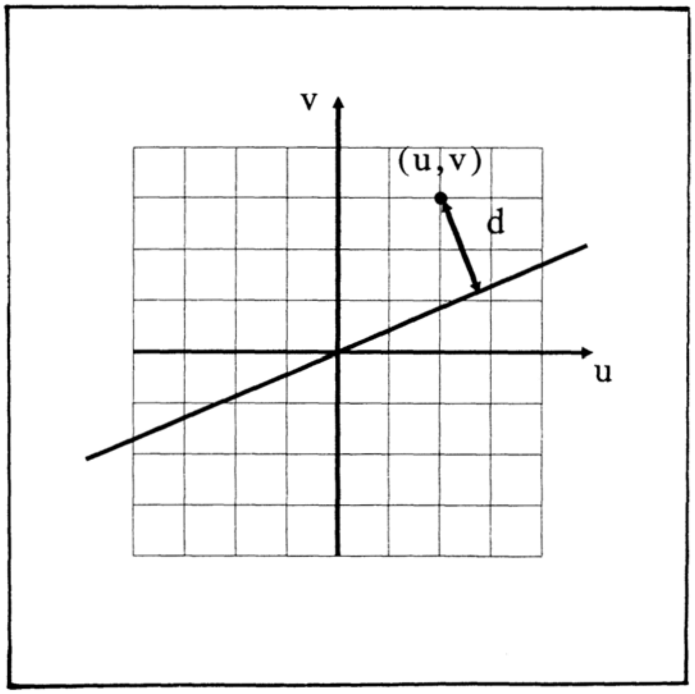


Figure 10.17 A weighted regression to estimate θ.

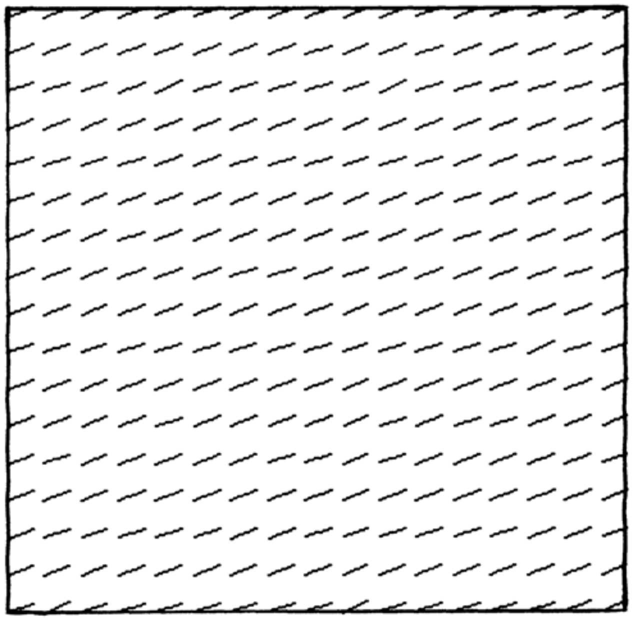


Figure 10.18 Results from test data at 70°.

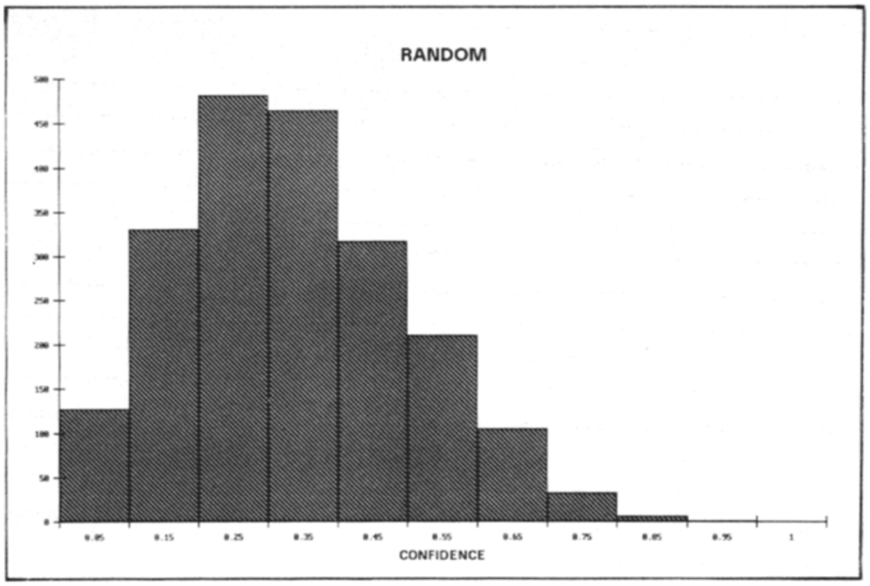


Figure 10.19 Random values of confidence.

The data set shown in Figure 10.13 was used as a test and produced the results shown in Figure 10.18.

In the diagram the angle computed for each point is represented by a short line, rather like a compass needle. This method of displaying the results has the added advantage that the *confidence* can be shown by either the length or the intensity of the little compass lines, making this essentially a four-dimensional picture. With this test data all the lines have high confidence as expected.

In order to interpret properly the confidence values which would be obtained, 3000 sets of 10×10 grids were generated so that each value was from a random rectangular distribution on (0, 10) using the Borland Turbo C generator rand () with a seed of 1. The confidence for each of these grids was then calculated and the results are shown in Figure 10.19.

Analysis shows that there is only a 10% chance of a confidence greater than 0.55, 5% chance of a confidence greater than 0. 61 and only a 1% chance of a confidence greater than 0.72, assuming that the data is random. These results are used in the next section to interpret confidences from real data.

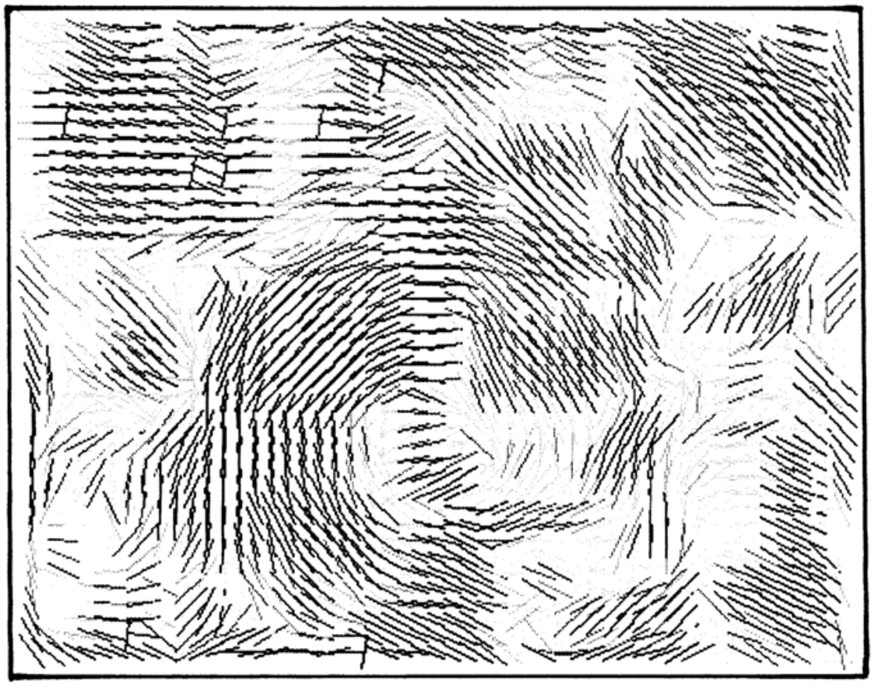


Figure 10.20 Ridge and Furrow at Stapeley Ring Cairn.

Results

Figure 10.20 shows the results when our technique is applied to the Stapeley data. The high intensity of the lines corresponds to high confidence. Again, the ring cairn is plainly visible. There are, however, several patches of high confidence, all with local uniformity of angles, and approximately similar angles between patches: those at top left are at about 100°, but those at bottom right are at about 120°. It is very interesting to compare specific areas in Figure 10.20 with Figure 10.11.

An analysis of both the angles and the confidences at Stapeley are shown in Figures 10.21 and 10.22; Figure 10.23 shows the distribution of angles at Stapeley for which the confidence is >0.5. These clearly show that an angle of between 100° and 120° is favoured.

Figures 10.24 and 10.25 show the results at Stapeley when only significant directions are emphasized. Highlighting in this manner is considered to be an important aid in determining the value of the results of the investigation.

It is possible to superimpose the compass needles on the topographical model itself. Colour Figure 10.26 shows the directions which the program found, the lengths of the lines indicating the confidence values. In addition to the lit surface, simple residuals have been used to modulate the hue, so that local height variations are emphasized.

It can be seen that the red, purple and white areas correspond with the tops of ridges, whereas the colours in Colour Figure 10.5 are shifted to the slopes of the ridges. Without going to extremes of visual bombardment, the information of Colour Figure 10.26 could be described as six- or even seven-dimensional, displaying the local ridge variations, their orientation and their confidence, and a lit three-dimensional surface.

Conclusion

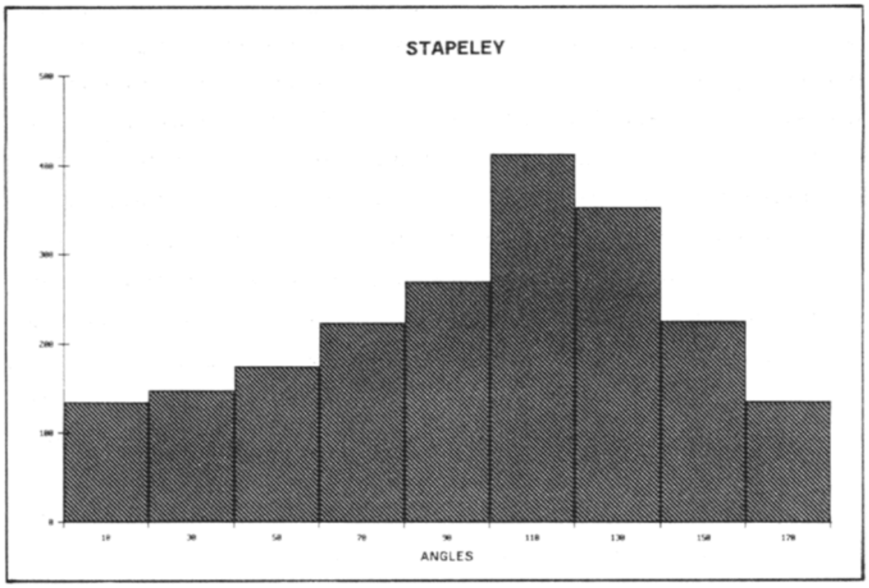


Figure 10.21 Angles at Stapeley.

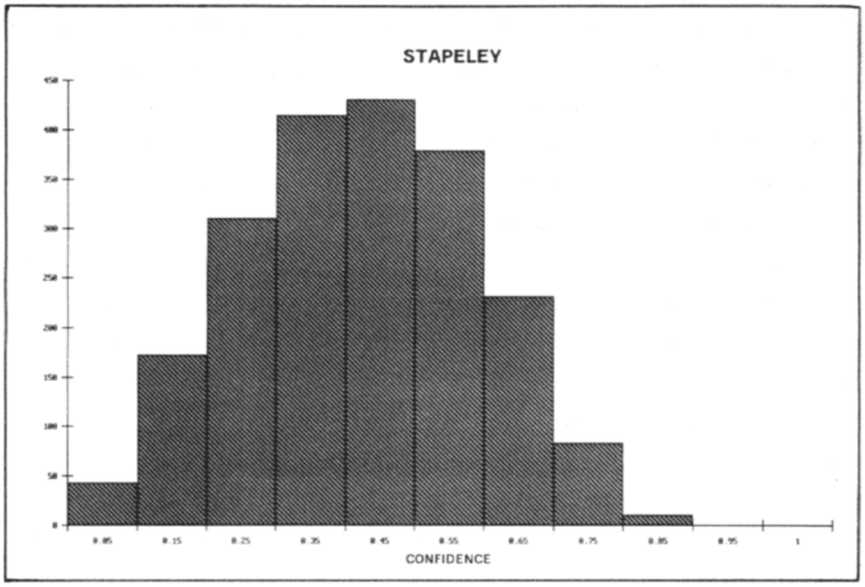


Figure 10.22 Confidences at Stapeley.

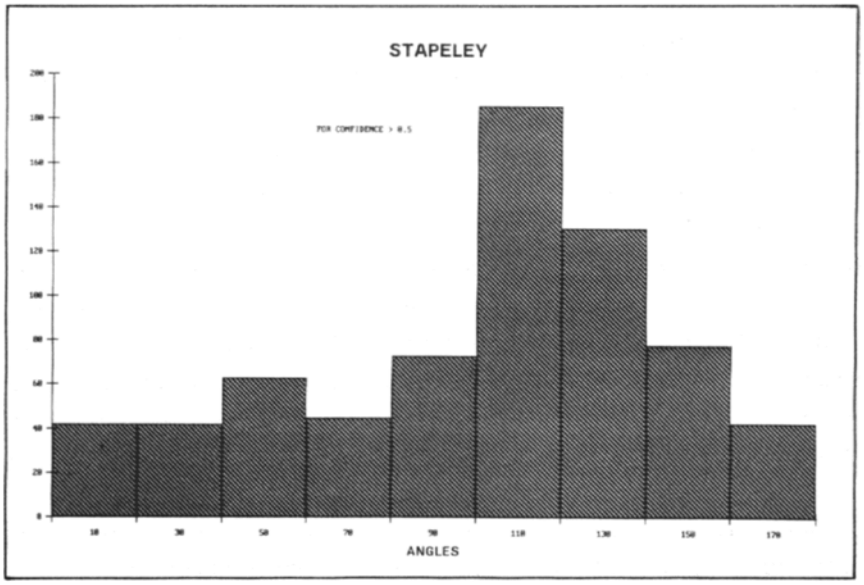


Figure 10.23 Stapeley for confidence >0.5.

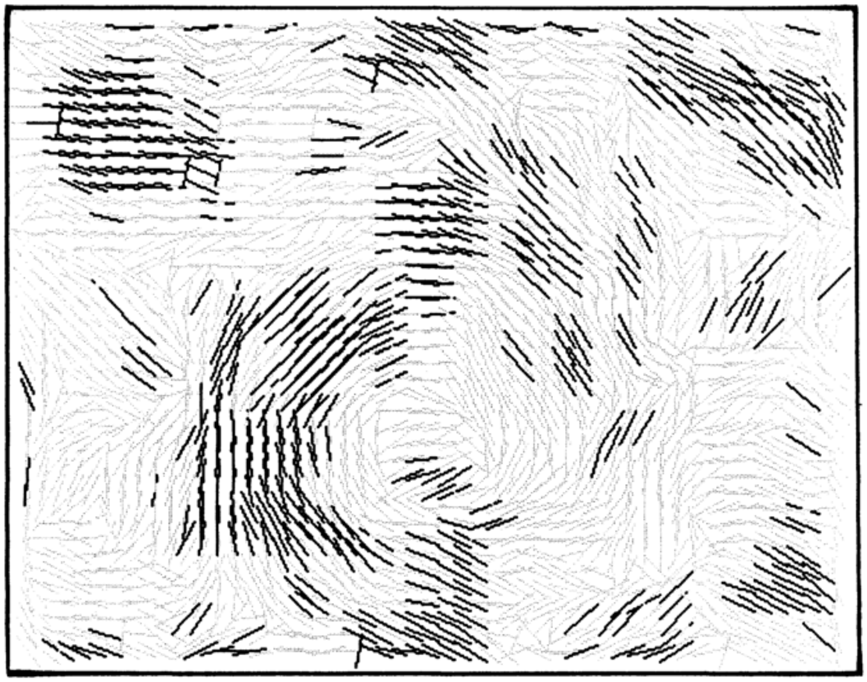


Figure 10.24 Significant results at 10% level.

This chapter has, we hope, demonstrated considerable success in objectively measuring the angle and confidence of ridge-and-furrow. It has also shown the value of appropriate choice of graphical display methods and the inclusion of confidence values as a fourth dimension. Furthermore, it has revealed the versatility of graphical techniques in highlighting other archaeological features, which should lead to more interesting results in the future. We believe that the ability to vary the display parameters in the manner of the examples shown here is an important investigative tool. We are currently taking these ideas further to estimate the pitch and phase of ridge-and-furrow and we expect then to deduce temporal relationships over the palimpsest. With the benefit of four- dimensional modelling we expect to be able to investigate further the possibility of a second mound or avenue at the site by undertaking geophysical surveys on the same grid as the topographical one, and displaying the combined results.

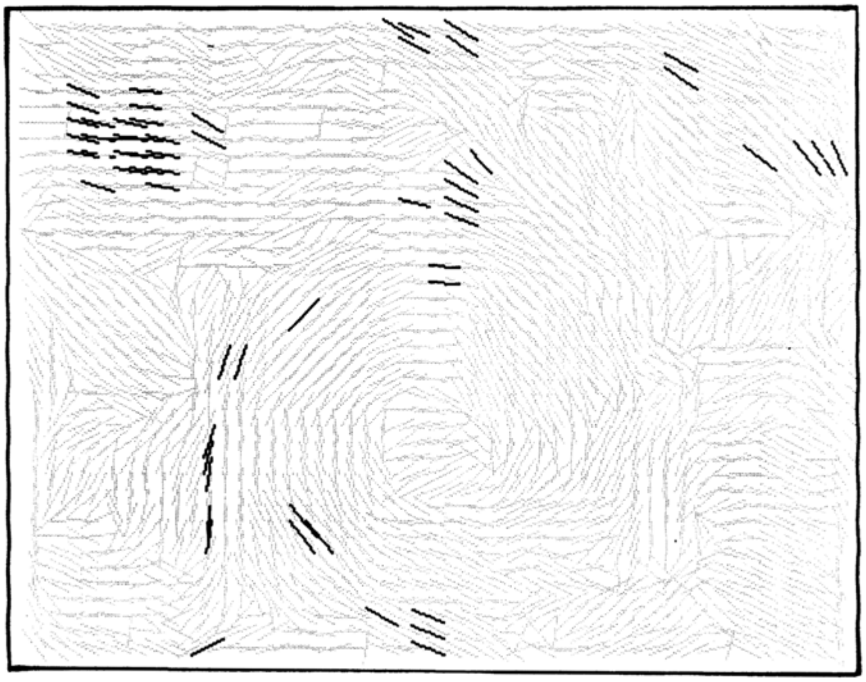


Figure 10.25 Significant results at 1% level.

References

Arnold, C.J., J.H.Huggett, P.Reilly & S.Springham 1989. Mathrafal: a case study in the application of computer graphics. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 147–156. Oxford: British Archaeological Reports.

Barker, P. 1986. *Understanding Archaeological Excavation.* London: Batsford.

Bracewell, R.N. 1984. The Fast Hartley Transform. *Proceedings of the IEEE* **72**.

Dougherty, E.R. & C.R.Giardina 1987. *Image Processing—Continuous to Discrete.* New Jersey: Princeton University Press.

Evans, J.G. 1972. *Land Snails in Archaeology.* London: Seminar Press.

Fleming, A. 1988. *The Dartmoor Reaves: investigating prehistoric land divisions.* London: Batsford.

Fletcher, M. & R.D.Spicer 1988a. Clonehenge: an experiment with gridded and non-gridded data. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 309–326. Oxford: British Archaeological Reports.

Fletcher, M. & R.D.Spicer 1988b. Clonehenge C-coded. *Archaeological Computing Newsletter* **16,** 4–10.

Fletcher, M. & R.D.Spicer 1988c. Experiments with gridded survey data. In *Computer and quantitative methods in archaeology 1987,* Ruggles, C.L.N. & S.Rahtz (eds), 201–220. Oxford: British Archaeological Reports.

Fletcher, M. & R.D.Spicer 1989. Stones and stripes. Unpublished paper. CAA Conference York.

Haigh, J.G.B. 1989. Rectification of aerial photographs by means of desk-top systems. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 111–118. Oxford: British Archaeological Reports.

Lock, G.R. 1980. Some archaeological uses of the PICASO computer graphics package. *Science and Archaeology* **22,** 16–24.

Reilly, P. 1988a. *Computer Analysis of an Archaeological Landscape: medieval land division on the Isle of Man.* Oxford: British Archaeological Reports.

Reilly, P. 1988b. *Data Visualisation: recent advances in the application of graphic systems to archaeology.* Winchester: IBM UK Scientific Centre.

Reilly, P. 1992. Three-dimensional modelling and primary archaeological data. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 12. London: Routledge.

Reynolds, P.J. 1980. The working agrospace of the iron age. *Landscape History* **2,** 1–18.

Richards, J.D. 1991. Terrain modelling, deposit survival and urban archaeology. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 175–82. Oxford: British Archaeological Reports.

Royal Commission on Ancient and Historic Monuments—Wales 1988. *An Inventory of the Ancient Monuments in Brecknock (Brycheiniog). The Prehistoric and Roman Monuments. Part II: hillforts and roman remains.* Cardiff: Royal Commission on Ancient and Historic Monuments (Wales).

Spicer, R.D. 1985. Stereoscopic representation of archaeological data—a case for drawing conclusions in depth. *Science and Archaeology* **27,** 13–24.

Spicer, R.D. 1988. Computer graphics and the perception of archaeological information: lies, damned statistics and…graphics! In *Computer and quantitative methods in archaeology 1987,* Ruggles, C.L.N. & S.Rahtz (eds), 187–200. Oxford: British Archaeological Reports.

Taylor, C. 1974. *Fieldwork in Medieval Archaeology*. London: Batford.

Taylor, C. 1975. *Fields in the English Landscape.* London: Dent.

Topping, P. 1989. Early cultivation in Northumberland and the Borders. *Proceedings of the Prehistoric Society* **55,** 161–179.

Wood, J. & Chapman, G. (with contributions by Ken Delooze & M.Trueman) 1992. Three-dimensional computer visualization of historic buildings—with particular reference to reconstruction modelling. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 11. London: Routledge.

Zucker, S.W. 1987. Early orientation selection: tangent fields and the dimensionality of their support. In *Readings in Computer Vision: readings in problems, principles, and paradigms.* Fischler, M.A. & O.Firschein (eds), 333–347. Los Altos: Morgan Kaufman.

11

Three-dimensional computer visualization of historic buildings —with particular reference to reconstruction modelling

JASON WOOD & GILL CHAPMAN (WITH CONTRIBUTIONS BY KEN DELOOZE & MICHAEL TRUEMAN)

Introduction

Until very recently, the cost of computer power and memory has put three-dimensional model-based graphical work beyond the reach of many archaeological units.[1] Due to the problem of funding (Blake 1989), most of the reconstruction modelling achieved so far has been as a result of contacts with specialists, often not archaeologists themselves, with access to the necessary hardware, software and operating expertise. This has inevitably meant that the software used has not always been chosen specifically with archaeology in mind, but has been appropriated to archaeology from a different primary purpose. It is also the case that the majority of this work has been generated on solid modelling packages of the kind used for engineering (Reilly 1992), rather than the surface modellers more commonly used in architecture.

Two case studies of work by the Lancaster University Archaeological Unit (LUAU) are introduced, to explore the relative merits of applying various modelling methods to the study and presentation of important historic buildings in north-west England.

Solid and surface modellers

The terms ‘solid’ and ‘surface’ used to describe three-dimensional computerized models, distinguish different methods of storing information about objects in the computer, resulting from different methods used to create them.

Solid modellers

The first type of software is the solid modeller. There are two main ways of representing solids—by constructive solid geometry (CSG), or by boundary representation (B rep). In a CSG model, the objects are made up of primitive geometric solids such as cubes, cylinders or spheres, which can be combined in various ways with the Boolean operations union, difference and intersection. In a B rep model, the solid is represented by its boundary faces, but the software holds information about the inside of the faces as well as the outside. In both cases,

calculations of physical properties such as volume can be carried out on the model. There are a number of projects where solid modelling techniques have been applied to three-dimensional reconstruction of historic buildings. A detailed discussion of these and other examples of solid modelling in archaeology is given by Reilly (1992).

Surface modellers

The second type of software is the surface modeller. The simplest way of displaying an object on a computer screen is as a mesh of lines (ie a wire-frame model). The only information which needs to be stored in the graphical database is the x, y and z co-ordinates of the vertices of each object, with a ‘pen up’ or ‘pen down’ code. The next level is to store information about the surfaces, or faces, of the object as well as its vertices. A flat surface will probably be made up of only one face, but a curved surface will be made up of several, and perhaps very many, flat faces. The information is kept in two files in the database. One lists the vertices with their co-ordinates; the other lists the faces of the object, with the vertices which are joined together to make up each face. This type of software is called a surface, or face, modeller. Hidden-surface removal is now possible, and the surfaces can be coloured and shaded in a variety of ways. Sophisticated surface modellers can produce very realistic images, but the models they produce are not true solids. The software only holds information about the surfaces of the objects; it has no means of calculating the volumes they enclose.

One of the first serious attempts in Britain to model an historic building in three-dimensional using surface modelling techniques was undertaken in 1985, as part of a comparative study by the Royal Commission on the Historical Monuments of England. A wire-frame model (with hidden lines removed) was produced of the large timber-framed medieval Governor’s

House in Newark (Nottinghamshire), using the RUCAPS software package (Fradgley, *pers. comm.*). More recently, a three-dimensional colour animation software known as Softimage has been used in conjunction with the Channel Four Television series ‘Down to Earth’ to model the neolithic enclosure at Avebury (Wiltshire), the roman villa at Stanwick (Northamptonshire) and the roman Basilica in London (Slight 1990). Other examples include the reconstruction of 13th century buildings at New Fresh Wharf in London, modelled on CAMEO 3D (Schofield, *pers. comm.*), and the modelling, direct from photogrammetric data, of the Tomb of Christ in Jerusalem (Israel), using Intergraph software (Biddle 1991).

Lancaster University archaeological unit

LUAU carries out a wide variety of archaeological work in north-west England and elsewhere. In particular, LUAU has considerable experience in the recording and analysis of standing historic buildings. Buildings surviving above ground, are just as important and fragile an archaeological resource as those remains buried below the surface. These days, the role of the archaeologist is seen as paramount in the production of detailed historic fabric surveys. Such analytical surveys are often required to achieve greater understanding of buildings in advance of, and during, major works of repair, conservation or alteration. Scaled plan and elevation drawings are the essential basis for detailed conservation proposals, and are often used for issuing instructions to building contractors.

LUAU’s recording techniques embrace all types of historic buildings, irrespective of date, function, material or state of preservation. The choice of survey methodology forms the basis of the recording scheme, and determines the potential for deriving an analysis and interpretation of the historic fabric. Dependent on the size and scale of the project, the basic data may include photogrammetric plots, rectified photography, or a combination of these with instrument-based control or hand measured survey. Different levels of recording are used in different circumstances. Levels range from full stone-by-stone recording of complex buildings, to selective recording of structures of more regular and repetitive build (Wood 1990).

Introducing the case studies

Building on the basis of a very considerable amount of recording and analysis which LUAU had achieved at Furness Abbey (Cumbria, UK), it was felt appropriate to extend the use of the survey data and to generate a three-dimensional model of a part of the monastic complex. To this end, a collaborative programme was established in 1988 with North Cheshire College’s Computer-Aided Engineering Centre (Warrington, UK) to investigate the possibilities for the application of three-dimensional computer solid modelling from which video sequences could be produced. North Cheshire College is one of only three educational establishments in the UK which holds a licence to use the very substantial and costly engineering software package known as PDMS (Plant Design Management System), developed by the CADCentre Ltd (Cambridge, UK). British Nuclear Fuels plc, also based in Warrington, are committed users of PDMS and exploiters of new technological developments such as EVS (Enhanced Visualization Software), a more sophisticated version of the CADCentre’s visualization module REVIEW. British Nuclear Fuels plc also have the technology to produce video from the PDMS database (Delooze & Wood 1991). It is interesting to note that a similar collaborative venture has been established by archaeologists and engineers working with computer systems developed for the nuclear industry in France. Here, software used by Electricité de France and known simply as CAO (Conception Assistée par Ordinateur) has been used to create an impressive model of the Temple of Amon at Karnak (Egypt) (Boccon-Gibod & Golvin 1990).

The second case study resulted from a collaboration between LUAU and Teesside Polytechnic’s School of Computing and Mathematics (Middlesbrough, UK). Archaeological survey data was requested by the School to form the basis of a post-graduate computer modelling dissertation. The site selected by LUAU for the project was that of the Langcliffe limeworks, near Settle (North Yorkshire, UK), in particular, the impressive remains of a Hoffman Kiln. It was considered that the kiln’s unusual shape would generate some interesting modelling problems. Surface modelling was undertaken using the CGAL (Computer Graphics Animation Language) software package, developed by Dr Peter Comninos at Teesside Polytechnic (Chapman 1990, pp. 44–70). CGAL had been successfully used to model the ‘capella ante portas’ at Furness Abbey as part of a previous student MSc project (Billington 1989).

The two reconstruction models produced for LUAU illustrate the differences between the two types of modelling programs. Neither system was chosen specifically for archaeology, and in fact neither was ideal, but nevertheless useful models have been produced in both cases, and important lessons have been learnt.

The purpose of the remainder of this chapter is to show how the programs were applied to the modelling of the two different historic buildings, and to assess the general archaeological usefulness and suitability of the softwares, and the implications for future work.

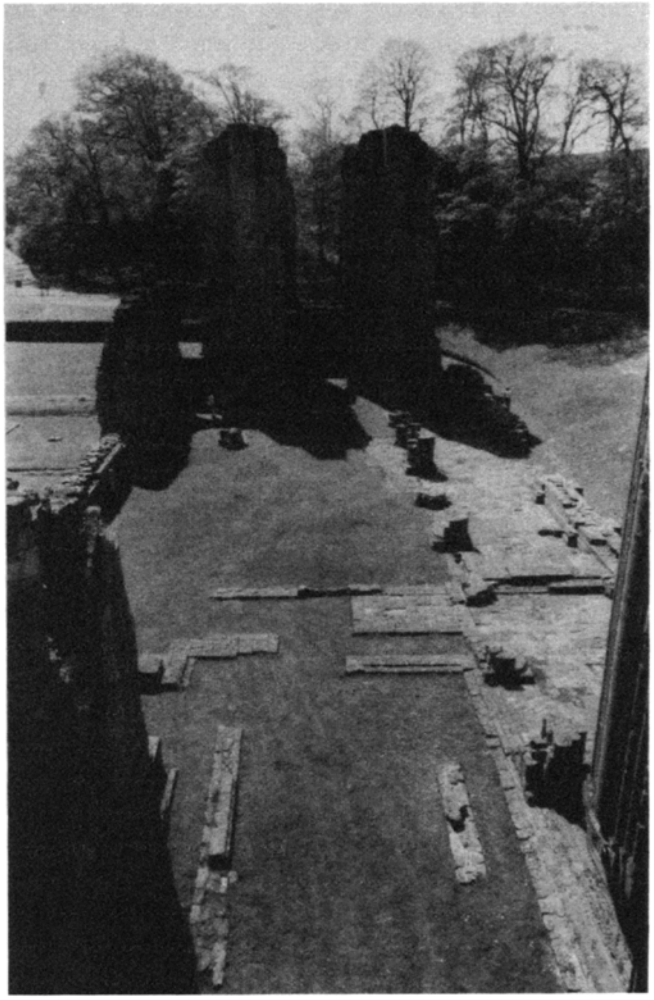


Figure 11.1 Furness Abbey—remains of the nave, aisles and west tower of the church, viewed from the crossing.

Case study one—Furness Abbey

Since 1985, LUAU has been co-ordinating a large scale archaeological and architectural survey and analysis of Furness Abbey. The site is one of the most substantial of the surviving ruined Cistercian houses in Britain, with the earliest remains dating from the second quarter of the 12th century AD. The project is one of a number of historic fabric surveys set up in recent years by English Heritage, to provide full basic recording of monuments prior to decisions regarding future repair and conservation programmes. The Furness Abbey survey was aimed at the production of full plan coverage and elevation drawings of the monastic buildings, together with analysis to assign periods and phases of construction, and, where possible, suggested reconstructions to each of the periods identified. The use of specialised survey techniques and methodical analyses has undoubtedly led to a closer understanding of the monument and its structural history (Wood 1987; Sherlock 1989; Wood 1991).

The abbey church

The area of the monument selected for modelling was that of the nave, aisles and west tower of the abbey church (Fig. 11.1). The nave and aisles were built in the latter half of the 12th century and were originally often bays, 50.50m long and 22.50m wide overall. The nave piers are alternately circular and clustered, while the aisles had simple quadripartite vaults, all of which have been destroyed. Virtually all of the nave arcade walling above the piers, and most of the aisle walls and internal screens, vanished during the Dissolution. At the end of the nave lies the west tower, built about AD 1500. The tower, with its large projecting buttresses, was constructed partially inside the existing church, thereby replacing the westernmost bay of the nave. Only the lower part of the tower, standing 17.75m high, now remains, the upper walling and heads of the eastern arch and western window having been destroyed (Dickinson 1965).

Some idea of the former appearance of the nave and aisle walls is gained from a study of the surviving transepts and crossing, which stand almost to their original height. The nature and spacing of the windows, as well as the position and angle of the former roof structure, are all evidenced in the extant fabric of the transepts and crossing.

Work on the abbey church involved a detailed survey and recording of the surviving structure, in both plan and elevation. Elevations were surveyed and plotted using photogrammetry, supplemented where appropriate by hand-measured recording. A combination of the photogrammetric plots and hand-drawn detail provided the information necessary to produce inked 1:50 stone-by-stone base drawings for each wall face. 1:100 colour-coded drawings were then produced to illustrate both the suggested historical development and reconstruction of the monument (Fig. 11.2).

Creating the model

PDMS was used to create a model of the abbey church. The software is used extensively worldwide by many large companies specialising in the design and construction of oil, gas, chemical, nuclear and power plant. As such, its main application area is that of plant layout, pipework design, drawing and materials control, and document management. The software also has interfaces to numerous related packages (such as stress analysis systems) which combine to make PDMS a powerful and indispensable tool for the design engineer (Pipes 1989).

The main benefit which comes from using the PDMS modelling system is that of integrity of design. The system's design database drives the drawing and materials control software, from which the fabrication drawings are produced. For a very large and complex model, there can be a great number of design databases which all must come together into one large single database for the complete project. The creation of only one 'intelligent' model, from which all design and construction data can be interrogated and extracted, leads to greater efficiency and consistency. As all drawings are taken from the single model, any design changes introduced will appear on all subsequent drawings. The system, therefore, allows for regular validation of the model's accuracy, and thus helps to reduce the number of fabrication and construction errors on site.

The main purpose to which PDMS is put is that of space management, which is visualized in three-dimensional views created within the design database. As well as pipework layout, the model also creates the building or room in which the plant is contained. It was, therefore, this architectural aspect of the software which was used to create the Furness Abbey model.

The program ran on a PRIME 9955, while the dimensional data for the model was supplied in the form of annotated two-dimensional reconstruction drawings, produced as part of the archaeological survey (Fig. 11.2). These drawings enabled the computer operator to accurately size all of the building's features. PDMS consists of many modules, accessed via a hierarchical software structure. For the abbey church model, only a relatively small number of modules were required:

1. Administration module—for setting up the project database; also a reconfiguror for database maintenance.
2. DESCON—the Design Constructor module.
3. GROCON and INTERCON—for grouping items and mirroring or copying.
4. GROUT—the Graphical Output module.

After a certain amount of exploratory work, it was decided to construct the nave and aisles on a stone-by-stone basis, with PDMS drawings produced using GROUT. To make the piers, for example, one stone was created, and then copied, and positioned until there were sufficient stones to form the first course. This course was then copied, the second positioned above it, and so on, until the whole pier was complete. Advantage was taken of using PDMS's special programs (or macros) for producing all of the various architectural features (piers, arches, windows, aisle responds, vaults, roofs, parapets etc.). Such 'elements' were duplicated by simply copying and mirroring the appropriate macros, thus saving much inputting time.

Once the initial PDMS model was completed, drawings could be produced at different levels of detail and at any desired scale. Multiple views, from any angle, with hidden lines removed could be made, either true perspective, wide-angle or conventional orthographic (Figs 11.3 and 11.4). In addition, sections, isometric and 'cut-away' views, and views from within the model looking out, were all possible.

Colour, lighting and model manipulation

The next stage in the model building process was to produce high quality colour-shaded pictures for multi-view walk-throughs. This was achieved using the EVS visualization module developed by British Nuclear Fuels plc. EVS runs on Silicon Graphics 4D series workstations and compatibles, and takes its model data directly from the PDMS database. In EVS, the PDMS elements are re-defined as polygons. The graphics buffer could hold approximately 60,000 polygons, but the first transfer from the PDMS model produced more than 60,000. It became necessary, therefore, to reduce the number of PDMS elements so as to allow the graphics buffer to manipulate the data at suitable speed. This was achieved by simplifying the

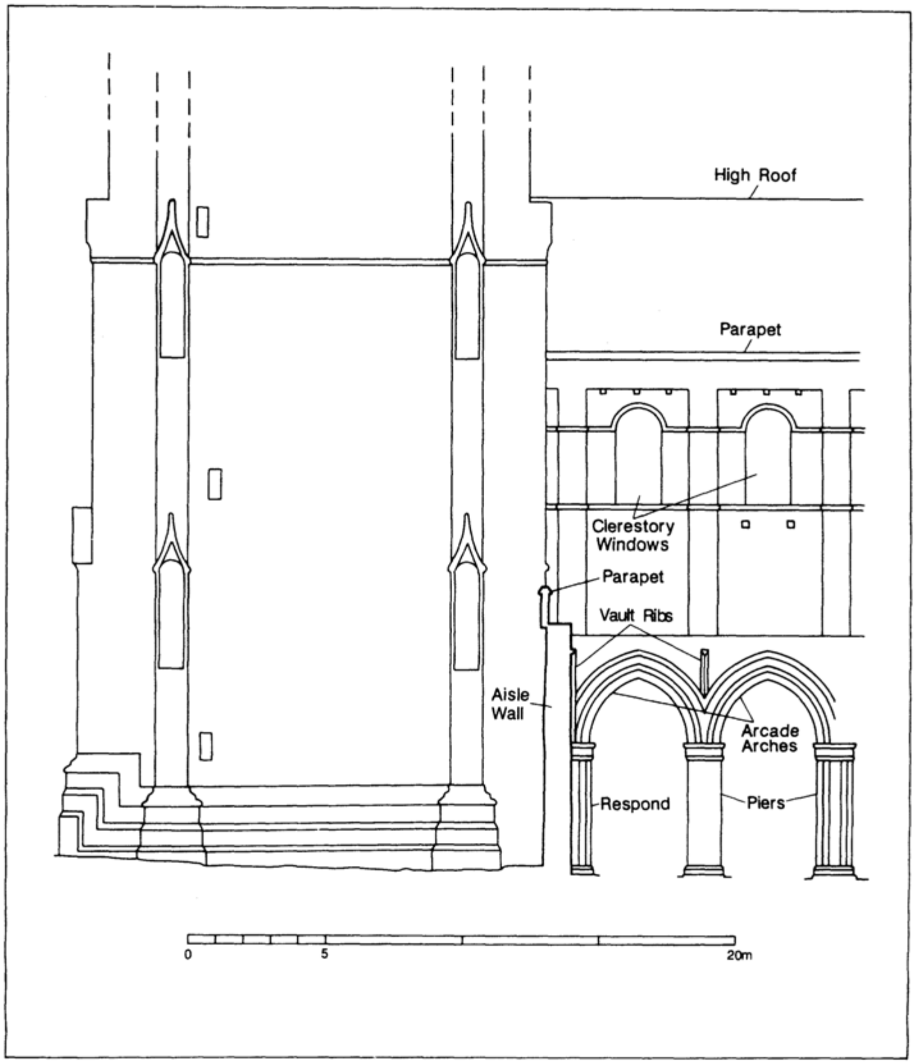


Figure 11.2 Furness Abbey—reconstruction drawing of the south faces of the west tower and southern nave arcade wall (bays 8 and 9).

PDMS model. For example, instead of constructing on a stone-by-stone basis, the piers were re-modelled using primitive cylinders; these were intersecting cylinders in the case of the clustered piers. The flat surfaces were also simplified; all the mortar joints within a wall face being modelled as one element. To complete this second version of the PDMS model, an undetailed representation of the west tower was added to the end of the nave. The second transfer into EVS resulted in the formation of 45,000 polygons. Although below the 60,000 ceiling, this number still proved too high for the machine to give good refresh times in quality mode. Real-time manipulation of the model was only possible by going into wire-frame representation and 'rapid travel' mode.

Quality colour-shaded pictures were viewed as both static images on the workstation and in hard copy as thermal transfer print-outs. The EVS module offered the ability to adopt any viewing position from inside and outside the abbey church, including both bird's eye and monk's eye views (Colour Figures 11.5 and 11.6). The ability to 'get inside and walk around' the reconstructed buildings gave a stronger feeling of enclosed space and volume, and enhanced the sense of 'being there'. The user-definable viewbox and zoom commands made it possible to select details from the model for particular inspection, and to create 'cut away' cross-sectional views through complex areas of the structure (Colour Figures 11.7 and 11.8). The ability to 'clip away' parts of the model was extremely useful for showing obscured detail such as the arrangement of trusses below the roof coverings and above the vaults (Colour Figure 11.5). The lighting control could also be varied to create the impression of depth and 'atmosphere' inside the reconstructed buildings.

EVS has no facility for generating shadows or reflections, and no texturing is available. The software is, therefore, limited in the context of creating a topographical landscape into which the abbey could be set. Such facilities are offered by other packages, but research and development would be necessary to interface with these.

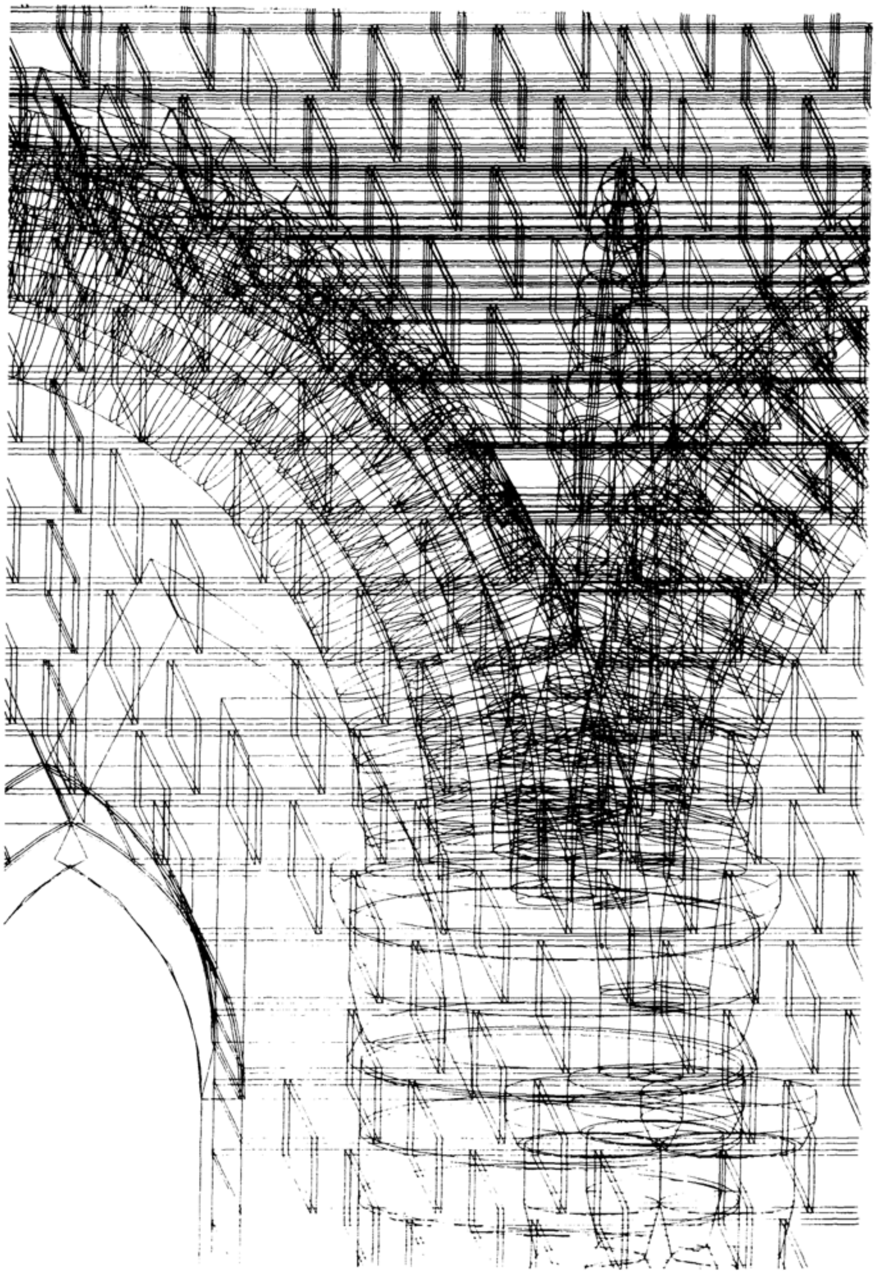


Figure 11.3 PDMS model—detail of GROUT plot of nave arcade with aisle walling and vaulting beyond, in wire-frame mode.

Animation

Following completion of the EVS model, an experimental video was produced using data taken directly from the PDMS database. The video was created by scanning (in 8 bits/pixel mode) the high resolution Red/Green/Blue (RGB) computer graphics display onto U-matic videotape via a RGB/Videolink. The simplified PDMS model, however, still proved to be very expensive in terms of computer storage, and led to problems and delays when the data was being made to generate a walk-through sequence. Real-time animation on videotape proved to be possible only by simplifying the model further.

The potential for the generation of real-time, colour-shaded animation sequences, direct from the original PDMS model, is currently being explored. Colleagues from the Warrington-based firm HELP (Hazard Evaluation and Loss Prevention) have reformatted PDMS models into a format acceptable to Amazing Array Production’s advanced WAVEFRONT graphics package. Software such as WAVEFRONT gives more control over every object, including the positions of camera and lights, producing a photo-realistic image (Pipes 1989).

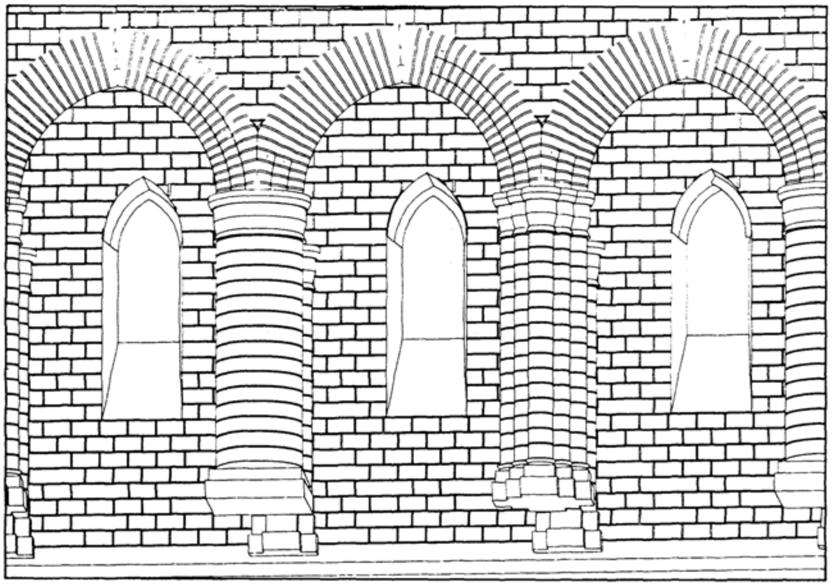


Figure 11.4 PDMS model—GROUT plot of nave arcade and aisle walling with hidden lines removed (compare with view in Figure 11.3).

Using the model

A computer model of this kind clearly has many benefits. For archaeologists dealing with historic buildings, the model can be used as a research tool to test reconstruction ideas, archaeological interpretation, and even, perhaps, simulation of structural stress (via appropriate interfaces). Attractive colour-shaded views also have a publicity and potential marketing value. Of greater potential, however, is the use of a model and video for presentation, education and communication purposes. A computer reconstruction would undoubtedly help to stimulate a greater interest and appreciation amongst the general public in the monument and its structural history. The chief fascination of the vast majority of people who visit monuments like Furness Abbey, appears to lie in a visual interpretation of the function and original appearance of standing buildings and excavated remains. A realistic picture is immediately recognisable to anyone, even if they know nothing about architecture or archaeology.

One possible theme for a video might be the depiction of how the site developed from earliest times, explaining how the various buildings were erected and in what chronological sequence. Another potential subject could be a surrogate walk-through video following a typical monk's day, showing what individual parts of the monastic complex looked like and were used for at a particular period. Views taken from the computer model could also be interposed with shots of the surviving ruins.

Further reconstructions of different parts of the abbey at different stages in its development are certainly possible. These models clearly have potential for depicting not only architectural features, but also the nature and position of temporary works such as wooden scaffolding and centring. Dynamic control of the colour palette and the ability to 'toggle' individual colours on and off would allow for differentiation between identifiable periods and phases of construction. Specific areas, such as floor surfaces, could be made to fade out and become transparent, in order to view buried foundations or earlier remains. Labelling of individual features within complex views is possible; the labels remain in place even while the model is rotated. Model components and labels can also be highlighted and made to flash against a duller background. Finally, an invaluable special feature of EVS is the interactive Scale Man or 'Monk'. This figure can be manipulated within the model to help provide a sense of scale and perspective, and can be moved independently of the model view. Of particular value is the ability to view the model 'through the eyes' of the Scale Man, by selecting the appropriate viewbox command.

The Furness Abbey case study explored the merits, problems and potential relating to the application of solid modelling techniques to the computer reconstruction of a complex and detailed historic building. The second case study, that of the Langcliffe Limekiln, was a smaller scale project. The study, undertaken on a much more modest budget, involved the application of surface modelling techniques, together with an assessment of the results which can be achieved in using such software to produce reconstruction models.

Case study two—the Langcliffe limekiln

In 1989, the Hoffman Lime Kiln and associated limeworks at Langcliffe were the subject of an archaeological survey in advance of proposed development of the site as a tourist attraction. The survey was carried out by LUAU on behalf of the

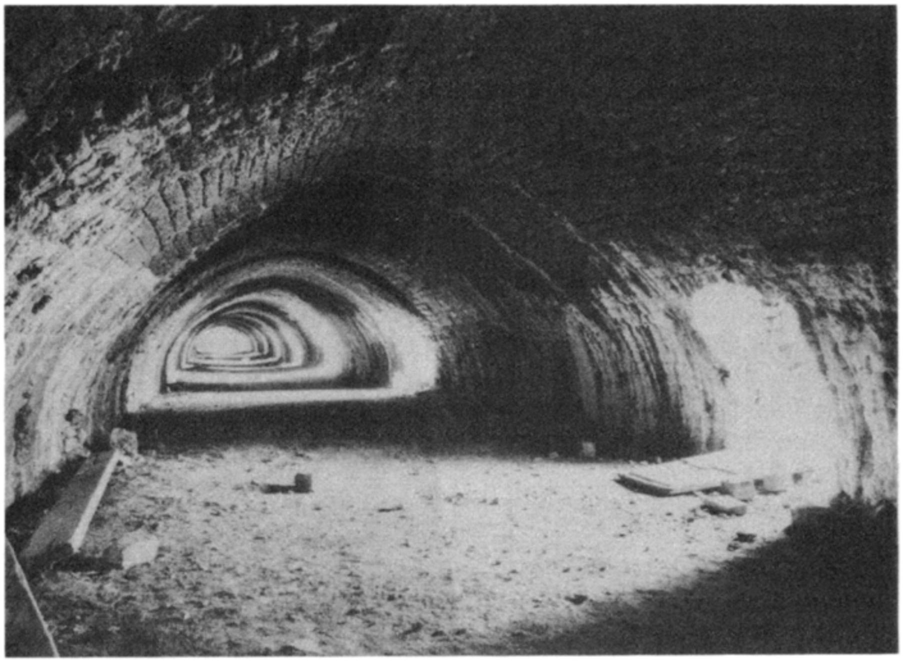


Figure 11.9 Langcliffe Hoffman Kiln—internal view of main tunnel.

Ribblesdale Trust and Yorkshire Dales National Park. The aim of the work was to provide a detailed inventory of the extant remains, and to assess the development and mode of operation of the limeworks using physical, documentary and oral evidence. The study provides a microcosm of the developments which affected the lime industry in the late 19th and early 20th centuries (Trueman, Isaac & Quartermaine 1989).

The use of lime grew rapidly during the Industrial Revolution and the Langcliffe quarries and works, sited next to the newly constructed southern section of the Settle and Carlisle Railway, were one product of this expansion. The original Hoffman kiln design was patented by Friedrich Hoffman in 1853 for brick making. Many improvements followed and the first use for lime burning presumably dates to this period. Hoffman himself, in 1870, patented a nearly rectangular version that became the more usual design. No methodical survey of Hoffman kilns in Britain has yet been attempted (Trueman 1990).

The Hoffman kiln

The area of the limeworks selected for modelling was that of the Hoffman kiln. The kiln, built in 1873, is an unusually large example of its kind, and survives today in a good state of preservation. The kiln structure, 128m long, 29m wide and 8m high, is rectangular in plan with semi-circular short ends. The external face of the kiln slopes steeply and is constructed of uncoursed, irregular limestone blocks. The interior consists of a continuous tunnel, 242m long, constructed of firebricks in the form of a barrel vault (Fig. 11.9). Access to this space is gained via a total of 22 barrel-vaulted tunnel entrances, cut through the thick outer wall of the kiln at regular intervals. A long, narrow, barrel-vaulted smoke chamber lies at the centre of the kiln, and into this run a total of 41 flues from the main tunnel.

The upper part of the kiln was originally surrounded by a curtain wall, which is now mostly in ruins. Virtually nothing remains of the once tall central brick-built chimney, or of the roof which originally covered the upper storey. A line of surviving iron rods indicates the positions of a series of dampers, used to control air flow from the flues into the smoke chamber. There is also evidence to suggest that a tramway once ran along the top of the kiln to convey coal, via a series of circular feed holes, to the limestone-filled chambers below. A water-balanced coal hoist, which facilitated delivery of the coal onto the upper surface of the kiln from adjacent railway sidings, was located at the southern end of the structure.

Work on the Hoffman kiln involved a detailed survey and recording of the surviving remains. Archaeological features were surveyed-in by Electronic Distance Meter (EDM) tacheometry, using a total station theodolite and data logger facility, and output to a portable plotter. Detail was then drawn by hand and a set of inked 1:200 plan and 1:100 cross-section drawings produced (Fig. 11.10).

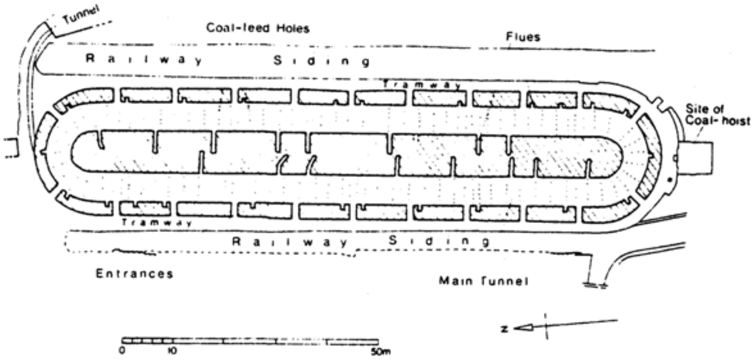


Figure 11.10 Langcliffe Hoffman Kiln—ground plan.

Creating the model

CGAL was used to create a model of the Hoffman kiln. The software is a surface modeller written primarily as a graphical animation package. The program ran on a networked system of Apollo Domain 4000 workstations, and colour-shaded views of the models were displayed on Ikon 64 bit graphics processors. Data for the model came from the archaeological survey in the form of the scale plans and cross-sections, together with explanatory text.

In CGAL, one works in a three-dimensional space where the origins of the x, y, z axes is zero. The user positions the objects and the light sources in relation to these axes. The kiln, for example, was positioned with the origin at its centre; the length of the kiln running along the x-axis and its width along the y-axis. Scale is defined in 'world units'. In the kiln's case, these represented metres. Once a model is built, it may be viewed from anywhere, and in any direction. These views are defined by reference to a viewing point and a centre point. Manipulation of objects (translation, rotation, scaling, etc.) and the defining of positions, is not interactive, but performed via keyboard commands.

The simplified model

The starting point for an object in a surface modeller like CGAL, is one or more two-dimensional shapes, or polygons. These can be generated in a variety of ways. For the kiln, x and y co-ordinates were taken from the archaeological plans and sections and entered via the keyboard. Once the two-dimensional shapes are complete, the software allows them to be manipulated in various ways to form three-dimensional objects. For example, a simplified kiln base was formed from two polygons: one the cross-section of half the base, including the interior tunnel and the upper curtain wall to its full height (Fig. 11.11); the other, the path taken by the outside edge of the kiln viewed in plan (Fig. 11.12). CGAL 'swept' the cross-section around the path to form the basic three-dimensional shape (Fig. 11.13). The simplified model was completed by the addition of the chimney and roof. The chimney was formed by 'lofting' a large square polygon and a smaller one. The 'lofting' process joins two different polygons together, providing they have the same number of vertices. The roof posts were made by 'extruding' single polygons. 'Extrusion' is the simplest method of manipulation performed in CGAL. The co-ordinates of each vertex in a polygon are simply copied and moved the desired distance in the third dimension, the corresponding vertices being joined to form new surfaces. In effect, this gives the polygon shape a thickness. The kiln was given a solid roof using the 'sweeping' method described above.

Viewed from above, the simplified model gave quite a good impression of what the kiln must have looked like (Fig. 11.14), but it was lacking in many respects, the most important of which was the tunnel entrances.

It is not possible in a surface modeller to cut a piece out of a model once it has been created, although some systems do now allow this process to be mimicked. Creating the kiln model with the tunnel entrances was to prove a much more complicated enterprise, and in fact turned out to be extremely difficult. For the simplified model, flat shapes were created and positioned on the sides of the kiln to represent the entrances. This was adequate for a distant view from above or from the side, but not for any close-ups or interior views.

The detailed model

The detailed model of the kiln took about ten times as long to build as the simple model. The creation of the final renderings and animations (see below) took at least as long again, but at this stage the computer was left running unattended for several hours, or sometimes days! The extra time taken to create the more detailed model was partly the result of more features being included in it, such as the coal hoist, the railway sidings and tramway tracks. The main problem, however, was the creation of the tunnel entrances and flues.

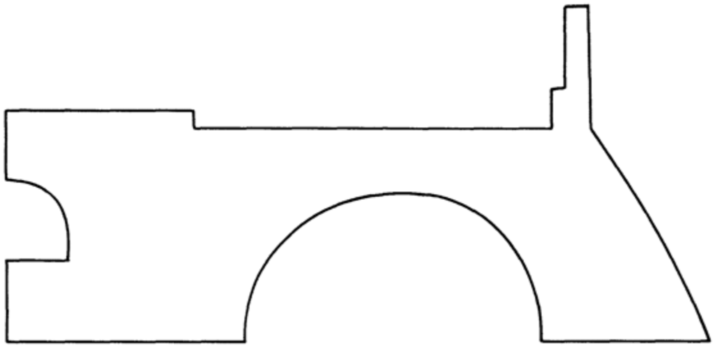


Figure 11.11 Simplified model—kiln cross-section.

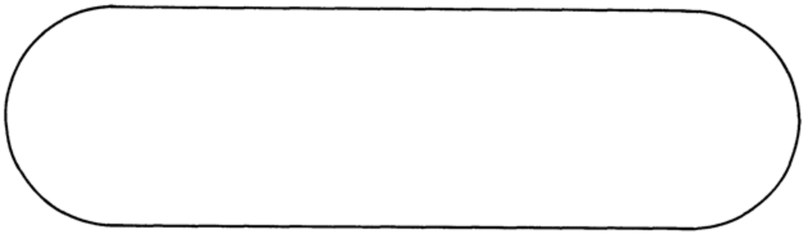


Figure 11.12 Simplified model—path for ‘sweep’.

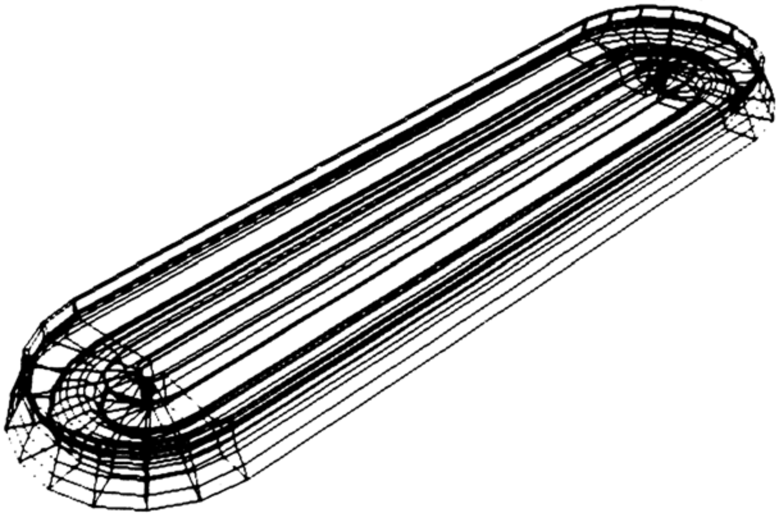


Figure 11.13 Simplified model—base of kiln.

Since holes could not be cut out of the model, it was necessary to build the kiln using lots of small segments fitted together. For example, a tunnel entrance was created from an inverted U-shaped polygon to fit the outside of the kiln, and another shorter one created for the inside. These could be rotated to fit the slope of the outer and inner walls respectively and then joined with a ‘lofting’ process to create the barrel-vaulted tunnel. At the shorter semi-circular ends of the kiln, the tunnel entrances fitted reasonably well to the outside, as the wall is only slightly curved at these points, but they did not fit at all well to the curving wall of the main interior tunnel. In order to make them fit better, vertices of the objects had to be moved— a fiddly and time-consuming process in CGAL because it is not interactive.

Once a tunnel entrance was completed, a segment of wall had to be created to fit over it. Several copies of the original polygon, representing the cross-section through the outer wall of the kiln, were made and altered, so that when joined together they would fit over the tunnel entrance. The joining method used here was a ‘skinning’ process. The polygons were positioned in space in an interactive subsystem of the program, and corresponding vertices joined to form a surface, or skin, over them. A separate block also had to be made to form the base of the section of kiln which contained the tunnel. Altogether, eight different ‘building blocks’ had to be made to allow for the tunnel entrances and flues, all of which had to be copied many times and positioned accurately. In this respect, the detailed model would have been simpler to make using a solid modeller, even though the initial kiln shape would have been more difficult.

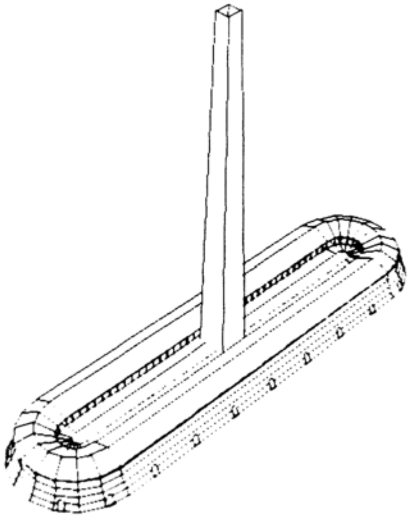


Figure 11.14 Simplified model—complete kiln with hidden lines removed.

The rest of the model was comparatively easy to create. The chimney, roof and its supporting posts were made in the same way as for the simplified model, except that the roof was made hollow for interior views. The rail track and tramways were made by ‘sweeping’ a small cross-section of track along a path, while the coal hoist was reconstructed from photographic evidence.

Colour and lighting

Wire-frame and hidden-line drawings are useful in many ways but it is when colour and light are added to the model that it really ‘comes alive’ as a three-dimensional object. These extra ingredients not only make the final rendered picture far more impressive than a line drawing, but they carry much more information as well; information about the three-dimensional shapes and the way in which they relate to each other. CGAL has a good colouring and lighting system. Colours can be specified in various ways, such as by palette selection based on the Munsell chart. Once chosen, the colours are applied to the different objects making up the model.

A good rendered view depends very largely on lighting effects. CGAL has up to twelve different light sources which can be placed anywhere in the model space. The computer then calculates how the light from each source affects each object and shades them according to the shading algorithm specified. It takes experience and some trial and error to obtain the desired lighting effects. Placing a light source where it would actually be in a scene may not give the desired result, not least because the light affects all surfaces which are turned towards it, even if there is an object in between! Solid objects do not ‘really’ exist in a surface modeller—light passes straight through them. As an example, the effect of light coming into the main tunnel through the entrances (Colour Figure 11.15) could not have been achieved with one strong light outside— this would have lit up the whole interior equally. Instead, light sources were placed in each tunnel entrance and the output from each controlled to ensure that the light grew dimmer as it got further from the source.

CGAL does not have ray-tracing, nor any other facility for rendering cast shadows or reflections, although these can be added using painting methods, given sufficient time and patience. The program does have the facility to alter the surface finish of an object—to make it more shiny, for example—and to apply various textures to it. These also take a considerable amount of trial and error to get right. The textures are three-dimensional in that they exist in space and show up on the object where its surfaces cut through them. Various different degrees of transparency can also be specified. However, the more facilities used, the longer the rendering takes.

For presentation purposes, some views of the model were taken into a computer painting package for the addition of detail (such as people to provide a sense of scale) which would be too time-consuming to model in three-dimensional (Colour Figures 11.16 and 11.17). This is not cheating, but a sensible use of different techniques for different purposes. Views of the model could also be used as a basis for hand rendering.

Animation

As well as providing discrete renderings, CGAL has reasonably good facilities for creating animations. The particular hardware available was not sufficiently powerful to enable real-time animations of anything but the very simplest of objects.

However, wireframe pictures could be stored and played back to check the movement. Paths could be created for objects to follow, or the viewing point could be sent along a path through the objects to create a 'fly-round' or a 'walk-through'. A small macro-program was written and the computer rendered and saved frames in batch mode. Twenty-five frames for each second of animation were saved; each frame for the kiln taking around twenty minutes to generate. Any kind of manipulation which can be performed in CGAL can be incorporated into the animation control program.

Once saved the frames were down-loaded onto a WORM (Write Once, Read Many) laser videodisc; another program was written to play the frames in the desired order, pausing on some if required. The finished animation was transferred onto VHS videotape.

Using the model

Like the abbey solid model, a computer model of this kind is useful both to archaeologists, for understanding the building and checking theories, and for communicating results to others. There are two major types of question which the model can be used to answer; the first being 'What did the kiln look like?'

Although the kiln remained in use until 1939, some details regarding its construction are still unclear. For instance, did the roof cover the whole upper storey, or was there an open space in the middle? The model can be used to jog memories. It can also be used to show what it might have been like to be inside the kiln during its operation —particularly appealing if the site is developed as a tourist attraction. Presentation views and surrogate tours can certainly contribute to this aspect.

Even more interesting in this case is the second question, 'How did it work?' Animation can enable the creation of a working model, showing, for example, the movement of air through the kiln.

Air came in through the tunnel entrances, passed through the fire, out through the flues to the smoke chamber, and thence up the chimney. An animation to illustrate this sequence showed the spaces inside the kiln as semi-solid objects within a transparent exterior. Coloured arrows were made to move through these spaces to represent air-flow (Colour Figure 11.18). Another useful animation illustrated the movement of limestone, lime and coal around the kiln. Again, the working model can be used both for theory testing and for communication.

Discussion and conclusions

This final section discusses the implications for data capture, an evaluation of the archaeological usefulness of both surface and solid modellers, and concludes with some recommendations on the suitability of software for the reconstruction of historic buildings.

Data capture

More research is needed to develop programs to allow integration of initial data capture on site, with the data format which is required for modelling. Three-dimensional co-ordinates relating to objects on site frequently exist in a computer database. These can be captured using traditional methods of measured survey, an EDM, or from photogrammetry. Theoretically it should be possible to use these co-ordinates directly for modelling. However, experience gained in the modelling of the 'capella ante portas' at Furness Abbey has shown that this is not generally practicable at present (Billington 1989, pp. 88–91). On this project, the use of stone-by-stone photogrammetric data to create a three-dimensional surface proved very time-consuming, due mainly to the incompatibility between the DXF file format in which the data arrived and the file format of the modeller, CGAL. A three-dimensional modelling package needs either geometric primitives or two-dimensional polygons as the basis for its objects. Polygons must have particular properties—for example, they must not be self-intersecting. Co-ordinate data taken from site rarely correspond to these criteria.

Facades generated from photogrammetric data can be 'extruded' (ie. given a thickness), and then positioned in three-dimensional to give a box which looks like the building from the outside; but there is no information about the inside of the box, and these are not three-dimensional models in the true sense. Moreover, the use of precise dimensional data taken directly from photogrammetry can introduce unwanted distortions into the model, particularly when the data relates, for example, to leaning walls. Another problem is that the level of detail recorded in this way produces far too much data for all but the most expensive three-dimensional systems to cope with, both in terms of disc space and speed of manipulation. This level of data is in any case unnecessary for modelling, the strengths of which lie in spatial relationships and only rarely in stone-by-stone representations or fine detailing.

Several things could happen to make it practicable to use this type of data directly for modelling. Those who record the data in the first place need to know what is involved in the modelling process so that they can organize their data sensibly from the start. Modelling software could perhaps become 'smarter' at understanding the modelling process itself, without the need for human intervention. The ever-reducing cost of computer power and memory will bring modelling in this kind of detail within reach of more users.

Evaluation of solid and surface modellers

Although most reconstruction modelling up to the present has been undertaken using solid modelling, the experience of the Langcliffe project shows that a surface modeller can also give useful results which may in fact be more appropriate to archaeology. Because the computer is holding a lot more information about a solid model than about a surface model, the application of realistic rendering techniques to it can use up an inordinate amount of computer memory, causing the use of such a model (such as the Furness example) to be both expensive and slow. If the main reason for making a model is to produce attractive views of a reconstructed building for public display, then the extra information contained in a solid model will be unnecessary and a surface modeller will give better value for money. An accurate surface model will also be useful to archaeologists themselves, enabling them to visualize the spaces involved and to communicate their ideas to each other. Although there are useful attributes to solid modelling programs, which cannot be duplicated in a surface model (‘slicing through’, for example), it is arguable whether the use of solid modellers justify the extra cost involved. There is the additional problem that until very recently, solid modellers have been extremely ‘user unfriendly’ and the construction techniques used do not deal well with complex and irregular shapes.

However, should archaeologists wish to put their models to more ‘scientific’ use, then a solid model might become necessary. Finite element analysis can be carried out on a solid model, enabling archaeologists to establish, for instance, whether building elements of a particular size and material could have stood up to the loads placed upon them (for example, the stress analysis systems which interface with PDMS). In the future, computer simulations, similar to those already used in medicine and meteorology, could be applied to reconstruction models of historic buildings. Simulation of climatic change, estimates of the quantities and placement of light, and the effects of decay, subsidence and collapse on a building’s fabric could perhaps be achieved. Models depicting different periods of construction (and a building undergoing construction), as well as visualizations of spaces and processes within buildings are all theoretically possible. Ideally, software is required which has the advantages of both solid and surface modellers, and these are now becoming available (see below).

Looking further ahead, the transfer of computer models and animation sequences onto interactive videodiscs would provide a useful and popular communication and education tool. Optical disc storage, in the form of the CD-ROM (Compact Disc —Read Only Memory), the WORM disc (as used in the making of the Langcliffe video), or more recently the CD-I (Compact Disc —Interactive), has sufficient capacity to store large quantities of graphical information. This can be entered into a database with other information and an interface written, so that members of the public can interrogate the data in an orderly but flexible way. This is the aim of the ‘Sacred Way Project’ (Cornforth, Davidson, Dallas & Lock 1990). In time, it may be possible for the public to access views of reconstruction models of buildings, and to ‘position themselves’ inside and outside them, at any given period during their history.

Suitability of software

Neither PDMS nor CGAL are really ideal for archaeological reconstruction of historic buildings, particularly if the aim is for archaeological units to start undertaking this work for themselves, rather than commissioning it from specialist centres. PDMS is far too ‘user-unfriendly’ for most archaeologists to contemplate, and in any case it runs on prohibitively expensive hardware. CGAL is not particularly ‘friendly’ either; its graphical facilities have been overtaken by newer software, and it cannot be interfaced with other proprietary software such as databases and draughting or painting packages. Despite these criticisms, both PDMS and CGAL have produced useful models, and much has been learnt from the rewarding and stimulating work which went into their creation. As far as the study of historic buildings is concerned, three-dimensional computer modelling can be a valuable exercise and is to be strongly recommended to archaeologists working in this field.

The first essential for those contemplating a choice of modelling software is to work out exactly what they want from their models. Attractive views for publicity? Accurate models to test theories and communicate with other professionals? Scientific working models? Interactive displays for education? Then they must decide what level of hardware they can afford. Three-dimensional graphics can begin on a good 386-based computer, while some units may be able to afford more powerful workstations. Other important criteria for choice include ‘user friendliness’, the variety of data input and output methods, standardization of file format, the level of support given, and the quality of available training (Chapman 1991b).

A survey of available software packages, undertaken on behalf of LUAU (Chapman 1990, pp. 71–115; Chapman 1991a), recommended the acquisition of Autocad 11 as the basis for the unit’s own computer modelling projects. Until recently, Autocad was not considered to be a good three-dimensional package, but its surface modelling capacity is much improved in the latest versions, and Release 11 has a solid modeller integrated to it. This allows a model created using surface modelling techniques to be converted into a solid model, ‘cut into’, and converted back into a surface model for rendering. Finite element analysis can also be carried out on the solid model if required. Autocad, although not the most ‘user-friendly’ package available, has the advantage that it is already widely used by archaeological units as a two-dimensional draughting package. Another suite of packages worthy of consideration is marketed by Intergraph. Their rendering and animation facilities are particularly good, and there is also a solid modelling option. Moreover, modelling can be carried out directly from photogrammetric data more easily than in Autocad.

There have clearly been some interesting and useful results from the work of a small number of practitioners. It is now both desirable and practicable that the benefits of three-dimensional computer modelling should become more widely available to archaeologists in general.

Acknowledgements

The Furness Abbey modelling project was funded by the Department of Trade and Industry, British Nuclear Fuels plc and English Heritage. Support has been welcomed from T.Keen, D. Pride and B.Sykes at North Cheshire College; J.Williams and A.Olivier at LUAU, and D. Sherlock at English Heritage. D.Cooper and R.Cooper of LUAU prepared the reconstruction drawings, based partly on photogrammetric data supplied by English Heritage (TSG, Survey Services). R.Busby of North Cheshire College and J.Durwood built the PDMS model, while the EVS pictures and video were created by K.Ryder and C.Williams of the Computer Applications Group at British Nuclear Fuels plc, under the direction of the late M.Hall-Wilton. Assistance was also provided by R.Longdon of the CADCentre Ltd and D.Grenfell of HELP. Travel grants to attend the Second World Archaeological Congress were gratefully received from Lancaster University, British Nuclear Fuels plc and the CADCentre Ltd.

The Langcliffe Limekiln modelling project was undertaken as part of an MSc dissertation at Teesside Polytechnic. Help and encouragement for the work was received from J. Webster, CAGTA Course Leader; S.Forgan of the Design Department, and S.Rutherford for photography and frame painting. The CGAL model was created from survey drawings prepared by S.Isaac and J.Quartermaine of LUAU. Permission to reproduce parts of the survey archive was granted by the Ribblesdale Trust. Assistance to the project in general was provided by R. White of the Yorkshire Dales National Park.

The authors are grateful for comments and assistance from a number of colleagues, and are particularly indebted to the editors of the present volume, Paul Reilly and Sebastian Rahtz, for their invitation to publish, valuable suggestions and improvements to the text, and general forbearance in awaiting the final version of this chapter.

Notes

1 This chapter is a much extended version of a paper which was presented in the ‘Communication in Archaeology: a global view of the impact of information technology’ symposia at the Second World Archaeological Congress, Barquisimeto, Venezuela, in September 1990. The original paper was concerned solely with the computerized modelling of Furness Abbey and is now more fully published (with illustrations) elsewhere (Delooze & Wood 1991). The section on the Langcliffe Limekiln model originally formed part of an MSc dissertation (Chapman 1990, pp. 44–70), and is here fully published for the first time. The section of the dissertation regarding surface modelling and proprietary software is reported on separately (Chapman 1991b; Chapman 1991a).

References

Biddle, M. 1991. Jerusalem: the tomb of Christ. *Current Archaeology* **11(3),** 107–112.

Billington, A.D. 1989. The representation of archaeological data by three-dimensional modelling techniques. MSc Computer Aided Graphical Technology Applications, Stage 3, Teesside Polytechnic.

Blake, V. 1989. A model way of working. *CADCAM International* **May,** 39–40.

Boccon-Gibod, H. & J.-C.Golvin 1990. Le Grand Temple d’Amon-Re à Karnak reconstruit par l’ordinateur. *Les Dossiers d’Archeologie* **153,** 8–19.

Chapman, G. 1990. 3D computer visualisation of archaeological data—with particular reference to reconstruction modelling. MSc Computer Aided Graphical Technology Applications, Stage 3, Teesside Polytechnic.

Chapman, G. 1991a. Do-it-Yourself reconstruction modelling. In *Computer Applications and Quantitative Methods in Archaeology 1991,* Lock, G. & Moffett, J. (eds). Oxford: British Archaeological Reports.

Chapman, G. 1991b. Surface modelling and proprietary software for building reconstruction. *Archaeological Computing Newsletter* **27**.

Cornforth, J., C.Davidson, C.Dallas & G.Lock 1990. Visualising Ancient Greece: the Sacred Way Project. In *Data Visualisation and Archaeology: visual reconstruction from partial data.* London: The British Computer Society.

Delooze, K. & J.Wood 1991. Furness Abbey Survey Project—the application of computer graphics and data visualisation to reconstruction modelling of an historic monument. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & Rahtz, S. (eds), 141–148. Oxford: British Archaeological Reports.

Dickinson, J.C. 1965. *Furness Abbey: Department of the Environment Official Handbook.* London: Her Majesty's Stationery Office.

Pipes, A. 1989. Walk-through plant designs. *Process Engineering* **229(6),** 12–13.

Reilly, P. 1992. Three-dimensional modelling and primary archaeological data. In *Archaeology and the Information Age,* Reilly, P. & Rahtz, S. (eds), chapter 12. London: Routledge.

Sherlock, D. 1989. Furness Abbey. *English Heritage Conservation Bulletin* **7,** 12–13.

Slight, D. 1990. Building a picture of the past. *The Guardian* **1st November**.

Trueman, M. 1990. A lime burning revolution. *Association for Industrial Archaeology Bulletin* **17 (2),** 1–2.

Trueman, M., S.Isaac & J.Quartermaine 1989. *The 'Langcliffe Quarry' limeworks, Settle: an archaeological survey of the site and Hoffman limekiln.* Lancaster: Lancaster University Archaeological Unit.

Wood, J. 1987. Furness Abbey, Cumbria: the initiation of an historic fabric survey. *Contrebis* **13,** 38–41.

Wood, J. 1990. Monument analysis. *Refurbishment Projects* **5(25),** 78.

Wood, J. 1991. Furness Abbey—recent archaeological work 1989–1990. *Archaeology North* **1,** 30–33.

12

Three-dimensional modelling and primary archaeological data

PAUL REILLY

Introduction

Data visualization and solid modelling are two terms currently experiencing increased circulation in archaeological circles. To many, the two terms are synonymous with advanced state-of-the-art computer graphics. Indeed, the picture generation and visualization aspects of these systems are their key properties. However, the considerable impact that many of these pictures achieve has detracted from other important attributes of systems that successfully structure and maintain complex interactions with enormous volumes of three-dimensional data. Comparatively few archaeologists have direct experience of these advanced systems and most do not fully appreciate the uses to which they can be put. For many, their only contact with this technology is through the generated pictures. They are then regarded as sophisticated presentation tools. While the pictures are indeed impressive, the modelling technology that makes such pictures possible is at least as important and interesting to archaeologists struggling with the problem of recording and interpreting primary archaeological formations.

What is data visualization?

The expression 'data visualization'—with its connotations of advanced computing, as opposed to non-computerized visualizations which may be traced back to Leonardo da Vinci and perhaps earlier—was placed firmly in the forefront of scientific jargon in the United States in 1987, after the publication of the National Science Foundation commissioned report entitled *Visualization in Scientific Computing* (McCormick, De-Fanti & Brown 1987). Data visualization is an evolving concept whose boundaries are continually expanding and, as such, is best defined in terms of loose generalizations. It refers to those techniques (generally the more technologically advanced) which allow visual interpretation of data through the representation, modelling and display of solids, surfaces, properties and animations, involving the use of graphics, image processing, computer vision and user interfaces. So it encompasses a much broader range of techniques than specific technologies such as solid modelling (see Wood & Chapman, (with contributions by Delooze & Trueman) 1992). The notion of data visualization is important to archaeologists and, more generally, scientists who work with huge volumes of complex data. Increasingly, we hear of supercomputing or numerically intensive computing (NIC), lengthy simulations and multi-dimensional data. These are manifestations of a growing need 'to see the unseeable'. Radar and other geophysical surveys provide numerous examples of this phenomenon of needing to visualize data and properties which are not directly perceivable by natural human sensors (eg. Nishimura & Mamei 1990). The difficulties of remote sensing underwater have stimulated some particularly excellent visualization projects (Stewart 1991). Data visualization is a means whereby much more multi-dimensional data can be brought within the range of human experience and cognition.

Archaeologists are not often the originators of totally new visualization tools, but they are extensive users of such tools and can have an important role in driving the development of novel approaches as users. Archaeological applications of solid modelling techniques are a case in point.

What is solid modelling?

Solid-modelling systems facilitate some of the most sophisticated computer graphics currently available and should not be confused with other graphics systems which produce pictures of 'solid-looking' objects. 'Face models', for instance, are composed of a set of rendered polygonal panels, which give the object the appearance of being solid but which do not actually conform to a truly enclosed solid shape. Solid modelling systems are intrinsically geared to the production of representations of solids and do not rely on the user to create the required set of faces. Today, most solid models are based on one of two data structures, and often both are present (Woodwark 1986). The 'boundary representation' (B-rep) model is similar to the face model, except that the faces, edges, and vertices of the model are linked together into a structure which is assured in its topological consistency. That is to say, there are no extra or omitted faces, edges, or vertices of the object. The 'set theoretic', or ‘constructive solid geometry’ (CSG) model, on the other hand, is defined as the combination of simple or primitive solids (eg. spheres and blocks), using operators derived from set theory (eg. union, difference and intersection).

A cynic might suggest that the original grounds for applying advanced modelling techniques to the reconstruction of ancient remains amounted to little more than an archaeological flirtation with an exotic technology in search of a new application, and resulted in mere intellectual curiosities. The case put forward here by contrast is that applications of solid modelling in archaeology are beginning to show real benefits in two disciplines which share a common goal in the definition and analysis of three-dimensional space. The view that the application of solid modelling to archaeology has much more to offer than just pretty pictures will be developed below.

It should be stressed that the graphical aspect of solid modelling systems is not necessarily their prime function. These systems also embody large volumes of structured three-dimensional information which can be exploited to establish links to many different kinds of database to provide powerful and flexible archival, training and analytic tools, and there is a great potential in archaeology to incorporate solid modelling within an integrated multi-media paradigm for the analysis and presentation of primary archaeological material. This potential grows if we take account of recent efforts to visualize free-form digital solids in connection with archaeological contexts and features. In relation to this, a further aim of this chapter is to suggest how the use of simulated three-dimensional archaeological formations can assist the development of new recording and analytical tools for understanding actual archaeological formations. Such facilities may play an important role in helping students to understand better the nature of archaeological features as well as being an excellent vehicle for demonstrating the relative benefits of applying different exploration scenarios (see Molyneaux 1992).

Historical overview of solid modelling in archaeology

B-rep models have only recently begun to be applied in archaeology, whereas set theoretic models have a moderate pedigree (see Chapman 1991 for a more detailed discussion of the relative merits of the two types of systems for the archaeological user). Since the mid-1980s, we have witnessed steady advances in the application of solid modelling techniques, particularly set-theoretic methods, to elucidate ancient monuments, with roman and romanesque building complexes being the most popular subject matter. These projects can be characterized as collaborative affairs in which archaeologists and computer scientists establish a symbiotic relationship. Of course, both parties usually have different perceptions and expectations of the benefits of such a relationship. The archaeological motivation behind the earliest projects was essentially to explore the potential of this sort of technology to illustrate monuments. By contrast, the computer scientists hoped to overcome certain technological problems in modelling complex objects.

There is little doubt that the earliest solid model reconstruction projects were considered successful and are still very influential. Certainly, once the models were built, they proved powerful and persuasive vehicles for illustrating the thoughts of archaeologists regarding reconstructions, especially of classical architecture. At the same time, the application of high-technology to the presentation of ancient civilization captured the imagination of the media (ie. press and television), which had the positive effect of publicising the wider projects, of which the modelling was just one aspect, to a much larger audience than might have otherwise been the case. As a bonus, archaeologists also realised that these models could stimulate useful insights relevant to the theory that certain social and political relationships are reflected in the use of space within the monument. This was demonstrated with the first computerized reconstruction of the temple precinct of roman Bath for example.

Woodwark successfully applied a set-theoretic solid-modelling system called DORA (Divided Object-Space Ray-Casting Algorithm, see Woodwark 1986) to the reconstruction of the temple precinct of roman Bath (Smith 1985; Woodwark 1991). Cunliffe supplied the initial data which consisted of a series of plans and elevations; dimensional data were provided for the temples of Sulis Minerva and the Four Seasons, a bath house and the precinct’s surrounding colonnade. All of this information was integrated into a computer model using a text input system called SID (Set Theoretic Input for Dora). In Dora objects are constructed using only one sort of primitive solid, that is the planar half-space (Woodwark 1991, p. 19.) Approximately 2,300 planar half-spaces were required for this model.

Once completed, Cunliffe examined the model from a number of different view points, through the medium of ‘view statements’ which acted as a ‘synthetic camera’. (The deciding factors on the limit to the number of views that could be computed were the power and availability of the computer.) Placing the viewer at the entrance of the reconstructed precinct suggested that a person standing in this part of the precinct would have been immediately impressed by the aspect of the temple of Sulis Minerva. In stark contrast, in the view from the top of the steps of the temple towards the entrance the attendant structures seem to shrink away. One response to this revelation is to ask whether feelings of superiority or inferiority were felt depending upon where in the precinct one happened to be standing? The model apparently can help sustain arguments relating to the precinct architects’ conception of the relationship between power and space. The power of solid models to convey certain types of abstract information is obvious, even from this one demonstration.

As well as opening up new possibilities of information dissemination, the use of this sort of technology can introduce limitations as well. Academic publishers are reluctant to increase the cover price of books by including colour plates, but the same colour images appeal to the producers of other media. For example, public interest in the Bath Temple Precinct project was heightened considerably in 1984 when the work on the model was featured in a television programme called *Waters of Bath* as part of the BBC's *Chronicle* series. This appeal to lay audiences is an important consideration to many projects, organizations and institutions. This is clearly the case in Bath, where the City Council has commissioned a second team of computer scientists from the University of Bath, in collaboration with local archaeologists, to build a computerized reconstruction of the civic baths of the roman town (Lavender, Wallis, Bowyer & Davenport 1990).

The roman civic bath complex of Bath city is one of the most visited monuments in England, with in the order of a million visitors each year (Colour Figure 12.1). The computerized model is intended to help visitors comprehend what they are looking at. The surviving remains of the roman baths of Bath, although indeed impressive, are only a faint echo of what was probably quite an imposing structure, and trying to visualize what the baths might have looked like is quite a problem for the uninformed lay person. First of all, the remains are, as the word implies, incomplete. Not having public access to all parts of what does remain exacerbates the problem. To make matters worse, a great deal of the remains available to the public are housed in the equally important, but in this case distracting, foundations of the georgian Pump Room. Since most people visit the roman baths only once, and their visit is limited to a relatively short period, the problem of 'instilling more than a modicum of superficial impressions' (Lavender, Wallis, Bowyer & Davenport 1990, p. 12) is a formidable one.

The solid model reconstruction of the baths, built using the DODO (Daughter Of DOra) program, is intended to illustrate how a roman citizen would use the baths by showing the various rooms in context, as they may have looked around AD 300. The model regards the baths as a 'sculpture' in stone, so that features such as water pools, doors and windows, and some of the mouldings are included in the model, but woodwork, pipework or painted decoration are not. Considerably more effort has been invested in modelling the interiors of the baths than was attempted in the precinct model. A measure of the success of the model in helping people to visualize the original structure is indicated by the fact that an early version of the model forced a rethink and redesign of parts of the reconstruction by the archaeologists involved in the project (Lavender, Wallis, Bowyer & Davenport 1990, p. 11).

On the technological side, the increasing amount of raw computer power available allows a larger number of views of the same model to be computed within shorter time frames. Probably of greater interest is the fact that more complex models can also be attempted. In this project, for instance, higher order surfaces, such as cylinders and tori, have been incorporated in the model. Roman columns are not true cylinders; in profile the edges rise from the base in parallel, but then taper off towards the upper end. This effect could have been modelled crudely using a truncated cone. Instead, the Bath team produced a more accurate representation by applying a technique for algebraically blending surfaces. (In fact, all the higher order polynomial surfaces in this project were built within the system by combining planar half-spaces algebraically.) The much larger number of primitives used in this model (12,000) is said to be near the limit of DODO's current capability (Lavender, Wallis, Bowyer & Davenport 1990, p. 9). No doubt we can expect significant improvements on this figure to be achieved in the future.

More powerful computers and increasing sophistication in software also enable greater realism to be introduced into the computed views of the model. Thus the rendered views of the bath models contained accurately modelled perspective, and so-called ray tracing techniques were employed to compute shadows. Ray tracing is one of the most popular rendering techniques, although so-called radiosity techniques are currently in the vanguard of computer graphics research. Briefly, ray tracing methods facilitate shading due to direct and indirect illumination and both reflection and refraction of light. Rays are cast from an 'eye' through the pixels on the screen into the three-dimensional space containing the model being rendered. When a ray intersects with an object in the model, calculations are performed to determine colour and other surface properties, such as gloss. At reflective or refractive surfaces additional ('daughter') rays obeying the laws of reflection or refraction are generated. The cumulative colour contribution of these rays is added to the pixel value to define the final colour.

Realism, however, is achieved not only by modelling natural lighting effects. Realism can be enhanced considerably by trying to embody some element of the social context and function of the subject in the visualization and thereby bring the visualization to life. Thus, in the first Bath project, a paint-box system was used by Bowyer to add figures in period costume to the computed views of the temple complex (see Colour Figure 12.2).

Additional ray traced effects were also introduced in the reconstruction of the legionary baths of roman Caerleon, another major solid modelling project involving an archaeological monument. The roman fortress at Caerleon was one of a string of strongholds built to control the activities of Celtic tribes living in the Welsh hills. Despite the dampness of the climate, some efforts were made to provide the sort of facilities which the roman army was used to in the Mediterranean and Gaul. The discovery, and subsequent excavations in the late 1970s, of the ruins of a large bath and gymnasium complex adjacent to the Caerleon fortress, located close to the present Welsh capital at Cardiff, comes as no surprise. During the 1980s, public access to this site was extensively improved, with the appointment of a full-time curator. Zienkiewicz provided Woodwark, and his colleagues in Bath, with data from which to build a reconstruction of part of the Caerleon complex, again using the Dora modeller (Smith 1985; Woodwark 1991). This turned out to be a much larger model than that of the Temple Precinct,

```
/* */ /* This routine generates a house with walls of a */ /* given thickness, centred along its z-axis. */ /*
Arguments passed to subroutine are: thickness, length, */ /* width, heightl, height2, rooftype, floor and rest. */ /
* Length and width give distance from centre of one wall */ /* to centre of opposite wall floor gives the
thickness */ /* of the floor in the house. If floor = then the */ /* building has no floor. WALLS and ROOF are
function */ /* calls and may be regarded as super-primitives */ /* */ dummy = (width+thickness)/2 do WALLS
(thickness,length,width,height1) COLOUR(colour-of-walls) UNION (ROOF
(thickness,length,width,height2,rooftype=‘hipped’, colour_of_walls=default_colour_of_wal; colour of
roof=default_colour_of_roof; overhang=default_roof_overhang ) AT(0-overhang,heightl,0)) /* for base of
house with standard shaped roof */ if fhˆ= 0 then do UNION(BLOCK (length+thickness,floor,width+thickness)
AT (0,0,0-dummy)) COLOUR(2) end end EXIT
```

Figure 12.4 Part of a typical WINSOM model file.

with some 7,000 planar half-spaces being used, and involving the modelling of a complicated roof structure and some internal detail such as a chessboard-like tiled floor. Moreover, as well as shadows, reflections and refractions were ray traced.

Animation was not feasible, but two coarse pseudo-animations were produced. One was a ‘flight’ around the complex, made as 36 views at intervals. The other was a ‘walk’ from the entrance to the compound, past a crudely modelled statue of Minerva and into the bath buildings, ending at one of the four stone basins known to be located in it. This sequence of some 120 pictures is housed in a permanent exhibit in the Caerleon museum; visitors look at the sequence through a colour monitor in the gallery.

The idea of creating multiple views to illustrate a monument was taken to its logical conclusion in 1985 when the first fully animated tour of an archaeological set theoretic model—that of the Old Minster of Winchester—was produced using the IBM UK Scientific Centre’s Winchester SOlid Modeller, WINSOM (Burridge, Collins, Galtin, Halbert, & Heywood 1989). WINSOM (now being superseded by WINSOM90) has a rich set of primitives, including planes, cubes, spheres, ellipsoids, helices and tori, and employs simple operators, such as union, difference and intersection to join, intersect or cut out shapes. Other operators are available to define the intensity and position of light sources, the viewer position and the type of ‘camera lens’ to be used.

Winchester’s saxon minster was demolished at the end of the eleventh century to make way for the present cathedral. A two-minute animation giving a ‘guided tour’ of the minster as it may have looked around the year AD 1000 was an integral part of the British Museum’s 1987 *Archaeology in Britain* exhibition. Several other versions of the animation appeared on a number of UK television broadcasts. The purpose of these animations was to depict the development of the minster over four centuries of royal patronage and to help people to appreciate the scale of construction as well as various spatial relationships between architectural elements within what was once one of the greatest buildings of medieval Europe (Reilly 1988b, pp. 28–36). Many of the experiences gained in the Old Minster project will be common to many reconstruction experiments using solid modelling methods (Colour Figure 12.3).

Kolbye-Biddle provided the initial data for the project in the form of detailed plans, elevations and isometric projections. Building the model was an arduous task, as model definition was through a text file input method. This meant that the programmers had to specify the dimensions of each primitive used, what operator should be applied to them and where they should be placed, in a special purpose programming language. A few examples will serve to illustrate what and how primitives were used: an infinite plane was used to represent the grass at ground-level; cylinders were used in pillars and arches; blocks (or cubes) were used for steps and walls; cones were used in the conical roofs on the apses; tori were used to add decoration to the altar pillars. The effort required to build the model was reduced by defining architectural elements which appeared in several places, such as windows, as self-contained subroutines. Differently sized windows could then be instantiated through a single call statement passing the appropriate parameters. The example code shown in Figure 12.4 is an extract from a larger program to model a building crudely.

Although detailed scale drawings were provided, positioning architectural elements still proved something of a black art. Early on it was appreciated that the careless use of an infinite cylinder, for instance, could wreak terrible havoc on the model.

Obviously, it was not sensible to produce high resolution shaded images each time the model builder wanted to check the latest change. Accordingly, other strategies were developed. One simple tactic was to generate lower resolution images. However, a much better and cheaper solution was not to shade the model at all, but to generate a wire-frame, using a WINSOM tool called FastDraw, which could be manipulated in real-time on a workstation (Burridge, Collins, Galtin, Halbert, & Heywood 1989, pp. 550–551). Both techniques proved invaluable in the building and verification of a reasonably large (c. 4,500 primitives) and complex model like that of the minster (Colour Figure 12.5).

FastDraw evaluates the intersection edges between the primitives and the spatial sub-division grid to produce the polygon sets.

The movie starts by showing the major building phases leading to the final eleventh-century minster. Once the minster model was finished, many hundreds of different views of it were computed to produce the surrogate tour.[1] This step from a single shaded image to a collection of related images which are presented in rapid succession introduced new problems. For example, choreographing the tour necessitated the writing of new programs to calculate the appropriate viewing parameters needed to give the illusion of smooth motion at each increment along the tour's route. In addition, at the time the *Minster Movie* was produced WINSOM produced 8-bit images, which meant that there were severe restrictions to the number of available colours and brightness levels (8 and 256 respectively), often resulting in unwanted banding effects. In order to produce smooth shading an error-diffusion technique was adopted. But while 'dithering' was very successful for removing the banding in any single image, it had the side-effect that corresponding pixel values calculated for adjacent frames might vary enough to produce a noticeable flicker. In fact a whole range of 'anti-aliasing' techniques had to be introduced to avoid these unsightly defects and produce an acceptable effect. Needless to say, the production of the minster movie involved what was a considerable investment in computing resources at the time (ie. long runs on a mainframe). Today 24-bit versions of the minster movie could be produced much faster using WINSOM90 on a workstation.

From the above commentary, it is apparent that the archaeologists involved in all the above projects had already formed fairly firm ideas about what the archaeological remains would have originally looked like before they considered building the computerized model. The solid models were really only a means of actualizing these ideas in a visually compelling way. The step of introducing the modelling at a much earlier stage in a project, and exploiting it as a malleable conceptual aid in the initial data interpretation process itself, was made during the investigation of the rather minimal remains of an early medieval Manx chapel and burial ground known as Malew 18 (Reilly 1988a, pp. 187–216).

The Malew 18 data consisted of a topographic survey which, when examined as a pseudo-illuminated surface showed positive features such as banks, a geophysical survey which, when image-processed, revealed buried negative features, such as ditches, and the barest possible record of an early twentieth-century antiquarian excavation of part of the site. The two surveys, combined, indicated the presence of a circular enclosure consisting of two concentric ditches with a bank separating them. The topography survey revealed vestiges of linear banks within the cemetery, and ditch-and-bank field boundaries appeared to butt onto the circular enclosure. In addition, the 'excavation' uncovered evidence of a chapel and graves.

Experiments to incorporate each of these features into a coherent WINSOM model brought to light flaws in the initial interpretations. What emerged was that, because the human modeller was forced to define explicitly the size and three-dimensional location of each feature in the model, inconsistencies were soon exposed. The result was that attention had to be continually directed back to details in the data. For instance, the complex was originally taken to comprise a continuous circular bank, with both external and internal ditches, yet, paradoxically, the ditch-and-dyke field boundaries butted onto the circular bank. The problem was that the bank appeared to float in mid-air at the point where the external ditch and the field boundary meet. While re-inspecting the geophysical data, particular attention was paid to some vague anomalies. These were now tentatively interpreted as causeways across the ditches. Besides speeding up the process of exposing and rectifying faulty thinking, the model building also had the effect of targeting key points in the complex where logical stratigraphic (ie. relative geo-chronological) relationships might be resolved. This is clearly essential information in order to understand the development of the complex.

Greater integration of the model-definition into the initial interpretation phase may also stimulate more opportunities to explore 'what if questions, enabling competing views to be compared and assessed iteratively.

Archaeological residues of the famous two-thousand year old temple site, within the iron age enclosed settlement at Navan in Amagh, Northern Ireland, also helped to extend and illustrate how solid modelling could be applied profitably to the analysis as well as the presentation of archaeological material. All that remained of the 'temple' was a series of concentric rings and avenues of post-holes. It appears that the structure was burnt down and deliberately buried under a cairn of stones. One theory suggested that the original structure may have been a wooden henge monument in which the position of the posts might be related to astronomic alignments. An opposing theory envisaged a large thatch-roofed building.

In order to throw some further light on this problem, a tentative CSG reconstruction of the temple was built to help the search for evidence of planning and measurement in its layout. In other words the model was intended as a tool with which to develop and assess different ideas and interpretations.

The models were developed using a rapid prototyping programming language called ESME(90) (Extensible Solid Modelling Editor) which has links to WINSOM90 (Burridge, Collins, Galtin, Halbert, & Heywood 1989, p. 550). Given a list of post-hole locations, the program computed the height of the posts, which would have been planted in the holes, on the basis of certain assumptions regarding a possible roof structure and then generated the appropriate model definition statements automatically. As the heights of individual posts were under the control of the ESME program, modifications due to even a minor alteration to the pitch of the roof, for instance, were straightforward to implement. If the refinements to the model were implemented manually, the degree of interaction with the model, in the building phase, would have been much lower.

The final model could then be fully rendered or displayed as a wire-frame. The interactivity of the wire-frame model enabled the archaeological investigator to explore the model for possible alignments. This project also featured an early experiment in texture usage in order to introduce some extra realism into the model. Traces of burnt reeds suggested the possibility of thatching. The effect of thatching on the cone-shaped roof was achieved using a novel texturing approach. The textures were not pictures mapped onto the surface of the model, they were solid three-dimensional patterns programmed using fractal geometry. The textured objects here are therefore analogous to sculptures carved in marble (Burridge, Collins, Galtin, Halbert, & Heywood 1989, p. 552; Todd 1990).

From the above discussion, we can characterize the main trends of the early phase of solid model applications to archaeology: first the projects have focused on reconstructing the upstanding elements of mostly monumental architecture. These pioneering projects have been very influential even though they have been largely regarded as publicity tools and technological advances have meant increasingly complex models are possible. Second, the shift in emphasis towards low level hypothesis testing of possible reconstructions as part of the model building stage of the Malew 18 and Navan projects has signalled a widening of the purpose of modelling; the solid model is no longer useful just as a pedagogic device, it has been adapted to become an analytical tool.

Recent trends and implications

These trends continue unabated. In Britain, for instance, the current modelling project of the Cistercian foundation of Furness Abbey is notable, not least, for the sheer size and complexity of the task. The surviving stonework at the abbey has been accurately measured, using photogrammetric survey techniques (Delooze & Wood 1990). A team from the Lancaster University Archaeological Unit and the North Cheshire College used a commercial plant design system called PDMS (Plant Design Management System) to build a model of the abbey, stone by stone, using data from the photogrammetric survey; each one of the many thousands of stones is an individual primitive separated by mortar primitives. A full account of the Furness Abbey modelling project is given in Wood & Chapman (1992).

From around 1988 onwards the number of teams developing solid model applications in archaeology started to grow noticeably. A number of different groups have begun to concentrate their efforts on particular aspects of the modelling process, most notably on user interfaces and photorealistic rendering. Another avenue being explored is that of exploiting the data visualization potential of these modellers at the much more fundamental stage of the data capture and analysis process.

The process by which the information defining the model is entered into the computer is a crucial one. If it is too hard or restrictive, archaeologists are unlikely to embrace the method and so human-computer interface problems in model definition and assembly are clearly important issues. Unwieldy-looking text file input methods are common, but are not attractive to many non-programmers. Consequently, more sophisticated, user-friendly methods are being sought.

A second modelling project whose subject is a Cistercian foundation—this time Kirkstall Abbey—is being carried out by archaeologists working with computer scientists at the University of Leeds (Colour Figure 12.6). Besides providing a solid model of parts of the abbey, the project is also trying to provide archaeologists with a user-friendly solid modelling tool-kit. The Leeds team believe that archaeologists want a ‘work bench’ on which to build models that do not require the user to spend time and effort acquiring a degree of competence in programming before they can build models. This dilemma was not encountered in some of the early projects because the archaeologist was completely detached from the modelling process itself; plans and elevation drawings were presented to the computer scientists, who were then expected to translate this into a model definition. Consequently, technology transfer into the archaeological world was extremely limited and the archaeological perception of the potential of the technology was restricted and, as a result, opportunities for developing the users’ requirements were surely missed. This project is also putting considerable research effort into realism. The rendering software presently being used allows for multiple light sources, shadows, reflection, refraction and three-dimensional solid textures (Dew, Fernando, Holliman, Lawler, Malhi, & Parkin 1990). The modeller employed here (Mistral-III) is particularly interesting because it is part of a research project exploring the opportunities for distributing computing tasks over an array of parallel processors (Holliman 1990).

Some interesting issues appear in the wake of this process of making modellers both easier to use and more widely available. For instance, the possibility of establishing libraries of architectural elements (or entire monuments) and consistent methods of model definition suggests that discussions on standardization are inevitable (Eiteljorg 1988).

Photorealism is a major research topic in the computer graphics world and it is not surprising to find that several computer scientists have chosen archaeology as a suitable application domain within which to follow their ideas through. Several archaeologists have also adopted modelling systems with very advanced rendering facilities. In fact, sophisticated textures are rapidly becoming required features in computerized archaeological reconstructions and consequently photorealism is one of the major goals in several interesting projects. This objective is the principal driving force behind the technological side of the work, stimulating both the development of novel algorithms and the push for ever-more powerful processing capability. Lighting and texture are currently the main topics being explored.

Probably the most notable recent example of the widespread use of texture, some incorporating permeable elements, appears in a three-minute animated tour of a model of Edo Castle, Tokyo. The Fujita Corporation together with IBM Japan has accurately reproduced the Grand Hall of Edo Castle and its ‘Pine-Painted’ walkway using Fujita’s COMPASS computer graphics system. The animation, rendered using the IBM Japan Tokyo Research Laboratory’s Rendering Subroutine Package (RSP), was broadcast as part of the *Kasuga-no-Tsubone* series on a channel of the Japanese public broadcasting station NHK (Nihon Hoso Kyokai). This beautifully detailed model illustrates Edo Castle as it probably appeared in Tokugawa period (1603–1867), the period of the Shoguns (see Nikkei Computer Graphics 1989a; Nikkei Computer Graphics 1989b; Miyata 1990 for graphics). The historical research behind the model was the work of Harai at the Tokyo Institute of Technology. Paintings decorating the sliding doors were reconstructed from documentary evidence. Ceilings, walls, sliding doors and openwork partition screens were reproduced by feeding images into the computer from a video camera and mapping them onto the relevant section of model; objects ‘behind’ partitions were rendered visible by incorporating permeation into the openwork elements of the images.

Ozawa has also been in the vanguard of solid modelling applications in Japan (Ozawa 1988; Oikawa 1992; Ozawa & Kawai 1991). His models include that of the fifth-century key-hole shaped mausoleum of the Emperor Ohjin and the settlement site of Yoshnigari, which was declared a national treasure in May 1991. In the visualization of the ancient village, a texture mapping technique has been introduced. The description of the basic structures of the houses is expressed in terms of logical AND/OR operations using planes and cylinders. In addition to ray tracing the model, texture maps were applied to the surfaces of the buildings to introduce more realism into the visualizations (Colour Figure 12.7). The models are based on data collected from many published reports of archaeological surveys (Ozawa, *pers. comm.*).

Another prime example of a project focusing on realism is provided by Cornforth & Davidson of Mechanical Intelligence who worked on a reconstruction of the N7 dining room and courtyard of the roman palace at Fishbourne in Sussex, England, to test and evaluate the performance of their modelling and rendering system (Cornforth & Davidson 1989; Cornforth, Davidson, Dallas & Lock 1990; Haggerty 1990). The textures used in this project are digitised pictures of mosaics and frescoes which are mapped onto the surfaces of the model. For the most part, the Fishbourne model is a B-rep model, but certain elements of the model were too difficult to model with this method and CSG techniques proved more expedient (Cornforth & Davidson 1989). The project relied on an early version of Pixar’s RenderMan package and was based on an experimental hardware configuration consisting of a microcomputer (Apple Macintosh II with 8Mb RAM) with key processes being accelerated using an additional parallel architecture (four Inmos T-800 transputers) in order that images be produced in acceptable times (Colour Figure 12.8).

The experience gained from the Fishbourne experiment is being built upon in a much more ambitious project to produce a surrogate tour of classical Eleusis, Greece; the tour will be available on DVI technology. The so-called ‘Sacred Way’ project is partially funded by the EEC COMET II initiative and involves two computer scientists, Cornforth and Davidson, collaborating with two archaeologists, Lock (Institute of Archaeology, University of Oxford) and Dallas (Benaki Museum, Athens). The planned modelling will incorporate the latest computer graphics techniques for realism, using both ray tracing and radiosity methods. The application of radiosity calculations in archaeological situations is a brand new proposal (Cornforth, Davidson, Dallas & Lock 1990; Dew, Fernando, Holliman, Lawler, Malhi, & Parkin 1990), but ray tracing remains the most widely used approach in modelling projects striving towards photorealistic output. Ray tracing is not considered entirely adequate for the Eleusis project because, although the results can be spectacular, the lights and shadows are too crisp. Moreover, the lighting is view dependent. It is hoped that by applying radiosity calculations as well as ray tracing methods much superior visualizations can be achieved (Cornforth, Davidson, Dallas & Lock 1990). Since radiosity calculations are currently some of the most computationally demanding another, more powerful, parallel processing architecture, based on the Inmos H1 transputer, will form the basis of the hardware.

The introduction of greater realism into the look of computerized reconstructions makes the models more attractive and interesting to a larger audience. It also raises some concerns. Most importantly, there is the issue of how best to inform the viewer on the degree of confidence that can be invested in each element of the reconstruction. Very often, for instance, no physical vestiges survive above the level of a building’s footings. A reconstruction may include architectural elements that have been inferred on the basis of evidence obtained from other comparable structures (‘parallels’) found elsewhere. Of course such models could be subject to structural analyses. However, many different stable structures might be built using a given ground plan, so other evidence is needed to lend authority to the chosen interpretation. The crux of the matter is that the models are largely interpretative; they are working hypotheses and are liable to change. It is not always obvious what the relation is between the recorded data and the interpretations which are built around them. There is a real danger of convincing the uncritical viewer that the model presented shows what the original really would have looked like. One way of dealing with this problem is to provide more than one model or a dynamic model. The final photorealistic interpretative model could be juxtaposed with a model of the raw data upon which the interpretation is based. Alternatively, the ‘control’ model would contain the same geometry as the interpretive model, but would incorporate some visual cues indicating the level of confidence associated with different elements of the interpretation. A visual method of denoting the level of confidence might take the form of colour-codes or variable levels of opacity. The problem might also be partially addressed by allowing researchers, and other interested parties, to link any part of an interpretation to the evidence on which it is based—a hypertext approach.

Relating the models to the underlying data

All the above projects are undoubtedly promoting the image of archaeology to the public and provide computer scientists with a fascinating test-bed domain. Another set of projects hopes to achieve these desirable effects while trying to harness the power of solid modelling to help deal with some fundamental problems in the collection and understanding of archaeological data and to find ways of enabling people to see how interpretations relate to the actual recorded material remains they are based upon. This has always been a critical problem for archaeologists.

There is a long tradition of recording archaeological contexts. Early accounts typically contain bald statements of the kind: ‘a roman floor was found’. Later, longer textual descriptions appear; the correctness and truth of the observation or interpretation being largely confirmed by the personal standing of the reporter (Hodder 1989; Trigger 1989, pp. 196–204). Gradually, the introduction of greater amounts of detail, such as scaled plans, elevations and perspective views, enable comparisons to be drawn. Photographs are also introduced. Photographs have the interesting property of being rather indiscriminate in the data they record, excepting, of course, that like the thematic illustration, they are static and view-dependent. Filmed sequences introduce a dynamic element, but remain constrained by the viewing angle, the lens used, the available light and so on. Although photographs are not always adequate to show subtle differences they are still regarded as important supporting evidence to interpretations.

Detailed records of formations are essential for many reasons. Not least are the issues raised by the destructive nature of excavations. At one extreme of the methodological spectrum, there are those who assess and explain a formation as they proceed, regarding their records simply as evidence for corroborating their interpretation of the formation. At the opposite end are those who try to excavate the site dispassionately, providing detailed objective records for later interpretation. Here the data are recorded as being independent of theory. Many other shades of opinion are also in circulation.

Generally speaking, every effort is made by excavators to record all relevant details concerning the nature of a deposit. The criteria defining what constitutes relevant detail are continually under review, and often there is a tendency to err on the side of caution, even when there is no clear idea about why certain details could be important. This is not to say that practitioners necessarily lean towards an empirical attitude to archaeological material. Nonetheless, an apparent pragmatic tendency towards the cautious approach, combined with an ever-increasing range of allied-and sub-disciplines who are interested in a wide range of different facets of buried archaeological formations, has led to a veritable explosion of data. Unfortunately, all recording methods (textual description, drawing and photography) are constrained by the limitations of the available technology, particularly the paper interface. Specifically, in projecting aspects of a three-dimensional space on to two-dimensional planes information is lost and the effectiveness of these tools is therefore circumscribed. A severe limitation is the strict view dependency. Nevertheless, these records are founded on a long tradition of convention and are well understood by trained and experienced practitioners. Scale drawings and black-and-white photographs also have the major attraction of being comparatively cheap to mass reproduce.

The nature of the limitations of a two-dimensional interface to three-dimensional data is very apparent in the archaeological sectional profile. While the excavation plan has the merit of having a direct bearing on some naturally occurring stratigraphic interface, either the top or the bottom of some archaeological context or other, the purpose of drawing sectional profiles of features, as a matter of procedure in the course of fieldwork, is much more difficult to understand.

The section can convey information on at least two important aspects of a context. First, it indicates the relative sequence of formations exposed in the section (eg. context X is below context W and cuts through context Y). A major shortcoming is that any number of intervening contexts may be involved, but do not happen to intersect with the plane of the section. Nowadays, information relating to stratigraphic order does not normally reside solely in section drawings, but is reduced to a series of logical relationships, usually maintained in Harris matrices. In addition to the limited stratigraphic information embodied in a section drawing, some shape information is also captured. Shape information is important in interpretation. Tips, dumps and truncated deposits have characteristic shapes for example. With no foreknowledge of the feature before investigation, decisions about where to place sections are frequently arbitrary in relation to the archaeological context. Therefore, if the feature is not symmetrical then a section must miss details. Clearly, better methods are required if this information is to be recorded.

It is not very surprising that the first computerized systems for recording and handling archaeological contexts should have inherited many of the characteristics of the traditional paper interfaces (Alvey 1989; Rains 1989; Stančič 1989; Weiss 1989, pp. 314–317; Alvey 1990). These first-generation computer-based site recording systems enhance considerably existing single context and other planimetric recording procedures (Harris 1989) by providing greater freedom to isolate and combine (ie. phase) contexts in smaller time frames, and their relevance and utility is undoubted. They can, moreover, be adapted to fulfil the role of supplying extra information about the shape of the contexts which the section drawing managed with only partial success. In principle, at least, any section can be generated from the planimetric data, providing that it is accompanied by sufficient three-dimensional surface readings.

As it will be shown below, preliminary experimentation indicates that more flexible records are feasible. In fact, developments in several technologies are creating a climate which could herald major improvements in what and how archaeological material is recorded, structured, analysed, presented and disseminated. These are hypertext, integrated multi-media systems, and three-dimensional modelling (including so-called virtual realities.) All embody techniques for representing and exploring data (Cornforth, Davidson, Dallas & Lock 1990; Loveria & Kinstler 1990).

A multi-media approach is in operation at Pompeii, the roman city buried by the eruption of the volcano Vesuvius in AD 79. As a result of a joint project between IBM Italy and FIAT Engineering (called 'Consorzio Neapolis') with the Soprintendenzia Archeologica di Pompeii, researchers have on-line access to the most complete set of photographs, maps, plans, sketches, archaeological reports, diaries, finds catalogues and computer reconstructions connected with the site which has ever been assembled, through workstations connected to a mainframe through a local area network (Colour Figure 12.9).

At Pompeii there has been a heavy reliance on graphics as an interface to the Pompeii archives. The most important navigation method through this colossal hypertext databank is via digitised maps of the city and its environs (Martin 1988; Gullini 1989; Moscati 1989; Zingarelli 1989). By clicking a cursor on part of a room in one of the buildings, scanned photographs of the room, or the frescoes on its walls, can be displayed. Help panels explain what the building was used for and how it was constructed. Technical words like 'hypocaust' are highlighted; by clicking on the word, a window containing a concise account of roman central heating systems appears. The user is prompted to look once more at the pictures of the room containing the heating system, to try and relate the explanation back to the actual building. Naturally, nothing can be seen of the heating system because, as the building is so well preserved, the system is still buried below the surface of the floors and walls. Here solid modelling comes to the rescue. Photographs can be replaced by corresponding views of a model of the same room, the only difference being that parts of the model are removed to reveal buried features including the hypocaust. Similar principles are being applied in other major Italian programmes such as the SITAG project on Sardinia ((Soprintendenza Archeologia Per le Provincie di Sassari e Nuoro) 1989, p. 31) (Colour Figure 12.10). The idea of integrating databases, surveys, reconstructions and other computerized data is also being explored by Australian researchers, working on the Syrian El Qitar project (Ogleby 1988) and by Americans working on material connected with the Peruvian Inka city of Machu Pichu (Chohfi 1990).

Impressive though such enormous projects are, a gap exists between the interpretation and the original data. It is not always readily apparent how one gets from the survey or excavation to the interpretation.

Virtual archaeology

Reconstructing archaeological sites is just one aspect of archaeological research. Understanding the subtleties of the raw data is, if anything, even more important to archaeologists themselves. By constructing detailed models of the excavated material, archaeologists can re-excavate the site and search for evidence which escaped attention during the actual excavation (at least to the tolerances imposed by the original investigative and recording methods used). Research of this kind clearly has major implications for how archaeological excavation and interpretation is taught as well as performed.

Experimentation with virtual archaeological formations may lead to new insights into data recording and analysis. The key concept here is 'virtual', an allusion to a model, a replica, the notion that something can act as a surrogate or replacement for an original. In other words, it refers to a description of an archaeological formation or to a simulated archaeological formation. (A simulated data set will normally be shaped by the criteria used for recording an actual formation). The problem is therefore to identify the quintessential components of the archaeological formation under investigation, since these must have implications for the styles of data representation and information handling that are possible.

In many cases archaeologists may not need to record in any greater detail than present-day standards demand. However, archaeologists must always pose and try to answer such questions as: 'to what level of detail can one record?' and 'at what level of detail must one record?' The overall archaeological objective of developing virtual environments must be to provide deeper insights into the understanding of archaeological formations by the addition of the powerful resources of the computer: a synergistic relationship. A number of studies are exploring the potential of simulation methods for establishing such a relationship.

The Southampton York Archaeological Simulation System project (SYASS) aims:

> to develop a simulation system to give students insights into the strategic decisions involved in planning and carrying out an archaeological excavation, with special reference to the costs of different strategies and the return in terms of different types of information. (Molyneaux 1992)

The original idea was to produce the archaeological equivalent of a flight simulator, not to simulate an archaeologist digging a site (O'Flaherty, Rahtz, Richards & Shennan 1990). The problem was therefore to encapsulate an archaeological excavation in a teaching program. Two key issues had to be confronted before this work could begin: 'what is excavation?' and 'why simulate?' The answer to the second question is rather more straightforward than to the first. In principle, simulated excavation is attractive because real excavation is destructive, expensive and slow; further, students do not get training in archaeological management and in any case most students will not become excavators. The answer to the question 'what is excavation?' involves many sub-questions such as 'what is a site?' and 'what are the questions being asked?' Most importantly here, there are issues of procedure: the benefits of utilizing particular techniques for non-destructive survey, research design, detailed recording, analysis and synthesis. The real question is therefore 'what are we simulating?' SYASS is not simulating an archaeological site, it is simulating what is known in Britain as Level 2 or 3 information.[2] To go beyond this requires a deeper analysis of archaeological excavation and its dependencies on the technology of excavation and recording, particularly those contingent on the limitations of the two-dimensional, paper-like, interface. This is not to decry the SYASS concept which is in any case intended to help students appreciate some of the decision making problems involved in fieldwork. What is being grasped at here is the possibility of simulating archaeological formations three-dimensionally and the development of new investigation procedures.

One problem brought out by Rahtz in his status report was that SYASS confronted students with concepts like 'context, spit, phase, horizon, locus' and so on before they had ever been on an excavation. To those not yet familiar with the finer points of trench credibility, such concepts can at best register as vague impressions on the mind or, even worse, meaningless jargon. What is required is a system to help users obtain a clearer idea about these entities without recourse to actual excavation. SYASS attempts to address this need through the use of non-schematic graphics on videodiscs. Thus the user can look at photographs of the context as well as the plans and sections. The SYASS concept is a natural successor to the simulation studies of Scollar (1969), Colley, Todd & Campling (1988) and Fletcher & Spicer (1989).

Scollar's original idea was to build a program for the simulation of magnetic anomalies of an archaeological origin as found using geophysical prospecting techniques. By creating an ideal set of values to which he could introduce noise, he began developing methods for removing characteristic distortions which could then be used to help in the analysis of real data sets.

Fletcher & Spicer developed a parallel idea for assessing archaeological topographic surveying methods. They created a mathematical model of an archaeological earthwork whose surface they could measure at any level of detail. They then proceeded to simulate a whole range of surveying methods, assessed them and made recommendations (see Fletcher & Spicer 1992).

Colley, Todd & Campling had the basis of an analogous system for excavation used in their Hamwic pit project (Colley, Todd & Campling 1988). Using the Winchester Graphics System (WGS) which coupled the power of a relational database system to a graphics sub-system, they analysed the distribution of the many thousands of objects they had found in an anglo-saxon rubbish pit. As the locations and positions of the objects deposited in the pit were recorded very accurately in three-dimensions the investigators were in a position to ask questions at many levels. For example, the huge quantities of bone could be studied as a complex three-dimensional point distribution. Alternatively, the spatial resolution of the distribution could be reduced to membership of a naturally occurring stratigraphic context, such as a layer, or an artificial spatial provenancing method could also be imposed. As most of the pit was recorded in very fine detail, it was possible to simulate recording methods of different precision to investigate the differences in results when applying different excavation strategies. The main limitation of this method was that, because data exploration was driven by database queries, questions relating to a spatial sub-set of the total assemblage were computationally inefficient and sometimes difficult to visualize. Moreover, it is extremely doubtful whether the lessons learned from exploring contrasting pit excavation scenarios could be applied with impunity in other archaeological circumstances. What is needed are comparable data sets for a range of archaeological formations: layers (eg. deep deposits, fine stratigraphy and tip layers), pits, recuts, post-holes, wells, walls, foundation trenches, artefact scatters, animal burrows etc. Attention is now being focused on modelling archaeological formations as they appear in the field. The challenge is no longer only to model buildings with simple geometry, but to model those amorphous humps, bumps and hollows, typically found in the course of fieldwork.

Solid models and archaeological contexts

Set theoretic solid modelling methods were introduced into the investigation of the early bronze age settlement site at Klinglberg-St.Veit in the Austrian Alps (Reilly & Shennan 1989). Normal methods of planning, levelling and sections through features, such as post-holes and pits, were used throughout the excavations. Nevertheless, attempts to build three-dimensional models of the deposits from the recorded data were in vain. Even though the excavators used the highest current standards of archaeological excavation, survey and recording, it could not be said that they produced a true three-dimensional record. The unavoidable conclusion is that most archaeological excavation recording has still a long way to go before excavators can claim that they record archaeological features in a manner that allows their full three-dimensional form to be reconstituted. The problem is that, at best, only the top surfaces of spatially extensive deposits are recorded suitably. This means that while it is possible to extract the three-dimensional shape of large layers, the shapes of cut features, which are reduced to an arbitrarily chosen sectional profile and a plan, are effectively lost.

This does not mean that no three-dimensional analyses are possible. One of the interesting problems with the Klinglberg data is the relationship of the material in the spatially extensive deep layers to that of intrusive features cut into the deposits immediately below them.

In order to make it possible to see whether material in the features was similar to and perhaps, therefore, connected with material in the local parts of the overlying layers, the digitised plans of the cut features (ie. pits, post-holes, etc.) were extruded to form prisms. These prisms were then intersected with a solid model of the overlying layers which, incidentally, were composed of box-shaped contexts intended to lend a degree of precision in spatial provenancing within these deposits. Colour-codes were employed to signify different properties (eg. average sherd size). The application of a clipping plane down the sides or to the top of the model could then be used to expose internal details and any visual correlations between properties of the cut features and local box-contexts could be assessed and, if interesting, investigated further.

This idea could be extended to help in the study of residual material if it were possible to model the cut features accurately. For example, it might show more clearly how some material in a deposit is derived from underlying contexts which have been penetrated by cutting features; the contents of the cut now forming part of the overlying material. Some steps towards building solid models of typical archaeological remains, as they are found in the field, took place in the Mathrafal project.

The site of Mathrafal in Wales has had a tradition of royal connections since the thirteenth century, when Welsh poets wrote in praise of the exploits of their leaders and named Mathrafal as their residence. The archaeological remains at Mathrafal comprise a series of impressive-looking earthworks, including a large mound and a rectangular ditch-and-bank enclosure (Colour Figure 12.11). While not under threat, there were several interesting questions concerning the phasing and function of the site which archaeologists studying medieval Welsh society wanted answered. Consequently, an integrated programme of non-destructive surveying techniques (ie. topographic, magnetometer and resistivity) was applied to identify the critical areas of the site where the minimum amount of excavation would yield the maximum amount of useful information. By mapping the geophysical survey data onto face models of the site's topography as colour-codes, the identified geophysical anomalies could be compared to local topographic features. As in the Malew 18 project the combined data sets showed some interesting details and a solid model reconstruction of several identified features, such as a building, kilns and a palisade, was built (Arnold, Huggett, Reilly & Springham 1989).

At this time, scalar fields, which could be handled with the same operators (ie. union, difference and intersection) as other WINSOM primitives (ie. plane, cube, sphere, cylinder cone, ellipsoid and torus), were developed to enable chemists to look at equipotential surfaces around molecules. A regular grid of three-dimensional values defined the location of field properties including membership of the set of points inside a complex shape (Burridge, Collins, Galtin, Halbert, & Heywood 1989, pp. 561–562). A variant of the fields primitive made it possible to integrate a solid model of the site's surface morphology with reconstructed components derived from the analyses of the non-destructive geophysical surveys (Arnold, Huggett, Reilly & Springham 1989). In combining the interpretation with the measured data, it is very easy to see how the two categories of information relate to one another. At the same time attention is redirected to unexplained features or anomalies which are left exposed.

This method of topographic solid modelling can, of course, also be extended down to the level of the buried context, but may require significant changes to the approaches and recording procedures applied in the field. Extensive deposits and cuts with clearly definable stratigraphic interfaces are fairly straightforward to characterize in solid modelling terms. Dealing with less tangible discontinuities, such gradual transitions through a formation, however, are more problematic. Nevertheless, the extra freedom that the field primitive introduces into the modelling arena has important implications for archaeological field recording methodologies, especially when we take into account parallel developments in other technologies which could conceivably converge in an integrated, seamless and multi-dimensional multi-media information environment.

A project known as Grafland tries to achieve this goal through an imaginary archaeological formation. Grafland is a simulated three-dimensional solid model of an archaeological formation containing layers, pits, post holes, cuts, recuts and so forth, and is a direct descendant of the data exploration and teaching experiments pioneered in research simulations like Clonehenge and educational initiatives like SYASS. It is intended to illustrate, in broad terms, how archaeological site-recording systems such as HINDSITE (Alvey 1990), might be extended through the use of a multi-dimensional representational tool, like a solid modeller, to be further enhanced by enabling links to a wider hypertextual system of the kind employed by the Consorzio Neapolis team.

Grafland's prime purpose is to demonstrate that archaeologists can produce more natural looking and intuitively obvious three-dimensional representations of buried archaeological data than has hitherto been practicable. Rather than reduce the record to a series of abstract single context plans and sections, each context is defined as a three-dimensional solid which can be examined from any elevation or sectional view. The 'naturalism' of the solid-modelled contexts is illusionary, of course, for archaeological excavators could never experience these entities in the manner that the solid modelling program allows, because the archaeologist never actually sees a whole or complete formation through excavation. Instead, archaeologists create and experience arbitrarily chosen archaeological veneers. At any given time during an excavation, part of the formation will have already been removed and part will remain out of view, buried beneath the exposed surface. It is by integrating each of these separate perceptions of the formation that a clearer idea of its totality may be reached. Recorded primary data is a metaphor for elements of the material continuum that is raw archaeological material. It is this metaphorical data that forms the basis of archaeological discussion. By enhancing and enriching the quality of this metaphorical data we hope to stimulate more and new archaeological discussions. It is at this juncture that the computer provides a powerful analytical aid in allowing virtual reconstruction of the primary data. Limited always by the quality of the recorded (metaphorical) data, the computer can be used to significantly enhance visualization, cognition and analysis of the structure of primary archaeological data. Unlike the raw material from which the recorded data was drawn, this metaphorical data can be dissected and explored repeatedly in an almost limitless number of ways. Individual contexts can be viewed in isolation or in free combinations. They can be enlarged, reduced or made transparent in order to explore some question. Moreover, contexts modelled as solid geometries are susceptible to a much wider range of transformations and interactions than the older methods of representation allowed. This increased freedom to explore multi-dimensional data sets opens the way for further insights into the nature of three-dimensional deposits and their recording.

The initial simulated excavation consists of a series of layers with various features cut into them. Most of the cut features—including a filled-in well, a palimpsest of intersecting pits with sub-contexts, beam slots, post holes, foundation trenches and an assemblage of artefacts—in the Grafland model are composed of compound CSG shapes, using the standard set of primitives (eg. cylinders and spheres) or parts thereof. Since the use of simple geometries may not provide a sufficiently convincing test bed for some archaeological practitioners, a more realistic model of an irregularly shaped feature, intended as a pit and containing artefact groups, has also been modelled using the WINSOM field primitive. Of course, other more complex models are possible, but the Grafland model is sufficiently complex for the purposes for which it was conceived.

The layers were manufactured by creating sets of hypothetical profiles, which were then digitised. This is equivalent to surveying the surface of each major layer using common transect lines. The logic of the process is exactly the same as removing and recording contexts in stratigraphic order. A layer (eg. context X) is defined initially as that volume between the measured surface and datum plane placed at some arbitrary depth below (call this object temp_x). Once this layer (ie. context X) is removed, the top surfaces of the context, or contexts, immediately underlying will be exposed. The tops of these contexts (eg. Y and Z) define the underside of the previously overlying context (ie. X). These contexts, in turn, will initially have default depth information (ie. temp_y and temp_z). The true geometry of a context is isolated using the standard CSG operators of union, difference and intersection (eg. X=temp_x difference (temp_y union temp_z)). Incidentally, the logical stratigraphic order of the deposits in the formation is embedded in the CSG model definition. The model could therefore be linked to a Harris matrix or phasing program, so that context sequences and connectivity can be studied. Knowing all the properties of Grafland, it becomes possible to devise different exploration scenarios to see how far they can facilitate a reconstruction of the site, the activities on the site and post-depositional processes operating at the site. Some of these capabilities exist piecemeal in several computer systems. What is important here is that they are all naturally integrated within the solid modelling paradigm (Colour Figure 12.12).

A Grafland animation sequence has been generated to illustrate the composition of the model excavation. The animation shows a flat green open space (a field) which gradually falls away leaving a block of ground (the simulated excavation volume) floating in space. This simulated formation is spun on its axis to show the different shapes of the major layers exposed in the profiles. Next, slices are cut away from one side, and later from the top, showing sections through pits, post holes and other cuts into the layers within the formation. In another sequence, each of the major layers is removed to reveal the contexts cut into the surface of the next major layer below. Each new layer surface is exposed in the order an archaeologist would meet them, that is the basic stratigraphic sequence. After this, the major layers are then removed from the visualization and only the cut features are displayed. The animation also shows these contexts being built up in reverse order and a zoom sequence into one feature passes through an artefact assemblage to illustrate the fine level of detail that can be recorded.

Such an animation brings out several key points. To begin with, the multiple views of the model demonstrate that the principle of constructing true three-dimensional solid models of archaeological formations is feasible and provides a superior record and database than orthodox methods currently allow, for further research. Allied to this, archaeologists can present larger quantities of complex data to a wider audience in easily intelligible forms. Archaeologists are therefore equipped to explain better how their interpretations relate to the data. Perhaps most important of all, data exploration and analysis are promoted still further. Visualization can be exploratory—in the sense that the researcher may pan through the data looking for loci of activity and other evidence. In other words, searches can be spatially organised, with the structure of the solid model being exploited as an efficient high-level spatial index. Conversely, the visualization can be more attribute directed. For example, if the modeller labels, or provides pointers to and from, component features, it is possible to isolate specific and associated stratigraphic components using standard database functions. An example would be a model in which all the cut features between layer and layer are isolated and displayed in order to study the different routes by which residual material could have travelled in getting from to . The solid model description has the additional benefit of having valuable quantitative details, such as the volumetric information about the contexts implicit in the model definition.

Conclusions and prospects

In the latter half of the 1980s and the early 1990s we have witnessed a steady progress in the application of solid modelling applications to archaeological problems. Initially the applications were implemented by computer scientists on behalf of archaeologists. This trend still persists, but the tendency is confined to projects attempting to exploit the latest technological advances in modelling or rendering. With the more established methods, we see a growing demand for archaeologists to be able to build the models. As archaeologists have become more sophisticated in the use of solid modelling systems, the scope of application has broadened from 'solid modelling to illustrate the monument' to 'solid modelling to analyse the monument' and most recently, 'solid modelling to relate the interpretation to the raw data' and 'virtual archaeologies'. Advances in modelling free-form solids mean that archaeologists can experiment with new recording strategies which supersede the traditional view dependent conventions. Solid models do not exclude the continued use of the well-established conventions; such schema could be extracted from the solid model definition.

Of course, modelling systems have also been advancing. Better interfaces mean that more people can usefully manage a modelling package, and enable further sophistication in the applications. For example, techniques of photorealism, such as transparency, may serve in 'solid models illustrating the hypothetical'. It was mentioned earlier, for instance, that transparency could be exploited to help the modeller convey the level of confidence associated with some element in a reconstruction.

The convergence of solid modelling and hypertext technologies opens up many interesting avenues which need to be explored in order to make the same archaeological record acceptable to those interested in preservation through recording, research, education and presentation.

Data visualization is the very attractive notion of representing data sets in graphical form so that they can be processed rapidly using the enormous cognitive power of the human brain at the huge band widths the human vision system is capable of sustaining. However, the power and impact of visualization is not simply a function of the visualization tool in terms of how much cpu is expended, or how well some observable phenomenon such as transparency can be modelled. The deciding factor determining the power of the visualization is how well it helps archaeologists to better understand and solve archaeologically relevant problems. Visualization on its own is not enough, because data has to be available in the correct format at the appropriate time. This implies that the visualization process should be an integral part of a larger data collection, archiving and processing environment. Fortunately, it seems then that the various technological and intellectual threads discussed in this chapter are coalescing.

A logical extension of the hypertext concept is to integrate solid models of the kind outlined in this chapter into a multimedia environment, not only as theoretical reconstructions, or even three-dimensional models of the recorded features, but also as user interfaces for data interrogation and navigation.

Hyperlinks may be introduced between the solid model and other data sets associated with the object of interest (eg. image, audio, video, DVI and text). A three-dimensional cursor could provide one possible interface, allowing users to 'point' at part of the model to discover what is being looked at and whether further information is available.

In the area of 'digital solids', in which free-form solids are modelled, we are witnessing exciting new developments. Already, modellers can extract feature data from sets of medical scans (eg. those produced in CAT) to build three-dimensional models of patients (Tyrell, Yazdy, Riley & Winterbottom 1990). Medical tomographic data is analogous to the geophysical scans produced from devices such as the 'Ground Pulse Radar', which is apparently capable of registering even small archaeological features many metres below the ground (Addyman & Stove 1989). However, there are two significant differences between the nature of the data embodied in medical and archaeo-geological scans, each of which represents a considerable challenge to routinely modelling and analysing archaeo-geophysical formations. First, the sheer volume of data is enormous and is already pushing hardware and software processing requirements. Second is the problem of feature recognition and extraction. Building models from scans of patients is made simpler because there already exists a considerable amount of *a priori* knowledge about the nature of human physiology.

At the moment, feature extraction is difficult with straightforward geometric models (Jared 1989). Looking for meaning in a virtual sea of heterogeneous three- (or more) dimensional data is one of the key problem-areas currently being addressed at the leading edge of the modelling world. However, there are many situations where non-destructive investigations would be a great boon for the profession. Developing these methods may help in delicate situations where, for instance, excavation might be regarded as profaning a sacred site. Equally, they should encourage us all to think more deeply about the physical nature of what it is we are investigating. Archaeologists should look forward to progress being made in multi-dimensional solid modelling with particular enthusiasm.

All this suggests that the principal impact of solid modelling on archaeology will be at a theoretically low level, with middle-level and middle-range theory perhaps experiencing secondary effects and high-level theories being unlikely to be affected at all.

In the meantime, Grafland-like models might be used as controlled data sets to devise and assess different excavation, recording and analysis scenarios. They may even prove helpful in evaluating the strengths and weaknesses of pattern recognition procedures.

Acknowledgements

Thanks to the following for allowing me to draw on their experience so freely: Richard Warner kindly provided the original Navan data. Professor John Woodwark provided the background information to the Temple Precinct of roman Bath project. Andrew Walter (IBM UK Scientific Centre) checked my review of the Old Minster of Winchester animation project. Professor Kazumasa Ozawa (Information Science Centre, Osaka Electro-Communication University) kindly sent me information relating to the imperial keyhole-shaped tombs and fishing village reconstructions. I am also indebted to Sebastian Rahtz for his comments on an earlier draft of this chapter. Discussions with Brian Molyneaux (Department of Archaeology, University of Southampton) have also influenced some of the arguments and comments put forward here.

Notes

1 Twenty five or thirty frames are needed for each second of the animation. In this instance, frames were ‘doubled-up’ so that only twelve different frames had to be computed instead.

2 In Britain level 2 information refers to the site archive and level 3 information is the archival interpretation. There are two other levels: level 1 is the physical data; level 4 is the synthesis.

References

Addyman, P.V. & C.Stove 1989. Ground Probing Impulse Radar: an experiment in remote sensing at York. *Antiquity* **63,** 337–342.

Alvey, B. 1989. Hindsite. *Archaeological Computing Newsletter* **19,** 4–5.

Alvey, B. 1990. Graphic interpretations of stratigraphy using desktop computers. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 1–6. Winchester: IBM UK Scientific Centre.

Arnold, C.J., J.H.Huggett, P.Reilly & S.Springham 1989. Mathrafal: a case study in the application of computer graphics. In *Computer Applications and Quantitative Methods in Archaeology 1989,* Rahtz, S. & J.Richards (eds), 147–156. Oxford: British Archaeological Reports.

Burridge, J., B.M.Collins, B.N.Galtin, A.R.Halbert, & T.R.Heywood 1989. The WINSOM solid modeller and its application to data visualisation. *IBM Systems Journal* **28 (4),** 548–578.

Chapman, G. 1991. Surface modelling and proprietary software for building reconstruction. *Archaeological Computing Newsletter* **27,** 1–7.

Chohfi, R.E. 1990. Integration of multidimensional data base and information system in archaeological research. In *Sciences Historiques, Sciences du Passé et Nouvelles Technologies d’Information: bilan et évaluation. Actes du Congrés International de Lille 1989,* Cacaly, S. & G. Losfeld (eds), 161–167. Lille: CREDO, Université de Gaulle, Lille III.

Colley, S., S.J.P.Todd & N.Campling 1988. 3-D computer graphics for archaeological data exploration: an example from saxon Southampton. *Journal of Archaeological Science* **15,** 99–106.

Cornforth, J. & C.Davidson 1989. Picturing the past. *Archaeological Computing Newsletter* **19,** 6–10.

Cornforth, J., C.Davidson, C.Dallas & G.Lock 1990. Visualising Ancient Greece: the Sacred Way Project. In *Data Visualisation and Archaeology: visual reconstruction from partial data.* London: The British Computer Society.

Delooze, K. & J.Wood 1990. Furness Abbey Survey Project—the application of computer graphics and data visualisation to reconstruction modelling of an historic monument. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 3: Late papers,* 1–8. Winchester: IBM UK Scientific Centre.

Dew, P.M., L.T.P.Fernando, N.S.Holliman, M.Lawler, R.Malhi, & D.Parkin 1990. Illuminating chapters in history: computer aided visualisation for archaeological reconstruction. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 15–25, & 117 (plates). Winchester: IBM UK Scientific Centre.

Eiteljorg, H. 1988. *Computer-Assisted Drafting and Design: new technologies for old problems.* Bryn Mawr: Center for the Study of Architecture.

Fletcher, M. & R.D.Spicer 1989. Stones and stripes. Unpublished paper: CAA Conference York.

Fletcher, M. & R.D.Spicer 1992. The display and analysis of ridge-and-furrow from topographically surveyed data. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 10. London: Routledge.

Gullini, G. 1989. Il computer per la rappresentazione del manufatto archeologico. *Rivista IBM* **XXV(1),** 42–44.

Haggerty, M. 1990. Remaking history one pixel at a time. *IEEE Computer Graphics & Applications* **July,** 3–6**.**

Harris, E.C. 1989. *Principles of Archaeological Stratigraphy.* London: Academic Press.

Hodder, I. 1989. Writing archaeology: site reports in context. *Antiquity* **63(239),** 268–274.

Holliman, N. 1990. Visualising solid models: an exercise in parallel programming. *Surface Noise* **3(September),** 11–15.

Jared, G.E.M. 1989. Recognising and using geometric features. In *Geometric Reasoning,* Woodwark, J.R. (ed.), 169–188. Oxford: Oxford University Press.

Lavender, D., A.Wallis, A.Bowyer & P.Davenport 1990. Solid modelling of roman Bath. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 7–13 & 116 (plates). Winchester: IBM UK Scientific Centre.

Loveria, G. & D.Kinstler 1990. Multimedia: DVI arrives. *Byte* **Fall,** 105–108.

Martin, J. 1988. Computing for art's sake. *Datamation* **December,** 86/9–86/13.

McCormick, B.H., T.A.DeFanti & M.D.Brown 1987. Visualisation in scientific computing. *Computer Graphics* **21(6)**.

Miyata, K. 1990. A method of generating stone wall patterns. *Computer Graphics,* **24(4),** 387–394.

Molyneaux, B.L. 1992. From virtuality to actuality: the archaeological site simulation environment. In *Archaelogy and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 19. London: Routledge.

Moscati, P. 1989. Archeologia e informatica: l' esperienza di Neapolis. *Rivista IBM* **XXV(1),** 24–27.

Nikkei Computer Graphics 1989a. Art and presentation. *Nikkei Computer Graphics* **12,** 74–79.

Nikkei Computer Graphics 1989b. Fujita industrial company and IBM Tokyo research laboratory work together to reproduce Edo castle as it was 400 years ago. *Nikkei Computer Graphics* **10,** 102–103.

Nishimura, Y. & H.Mamei 1990. A study of the application of geophysical prospection. In *Archaeolometry '90,* 757–765. Basel: Birkhäser Verlag.

O'Flaherty, B., S.Rahtz, J.Richards & S.Shennan 1990. The development of computer-based resources for teaching archaeology. In *Sciences Historiques, Sciences du Passé et Nouvelles Technologies d'Information: bilan et évaluation. Actes du Congrés International de Lille 1989,* Cacaly, S. & G.Losfeld (eds). Lille: CREDO, Universite de Gaulle, Lille III.

Ogleby, C.L. 1988. The Surveys of El Qitar: a late bronze age fortified city in Syria. In *30th Australian Survey Congress,* 321–336. Sydney: Australian Institute of Surveyors.

Oikawa, A. 1992. Japanese archaeological site databases and data visualization. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 6. London: Routledge.

Ozawa, K. 1988. *Zempokeonfun no suri (mathematical approaches to keyhole-shaped tombs)* Tokyo: Yuzankaku.

Ozawa, K. & T.Kawai 1991. Reconstruction of the Yoshinogari site using three-dimensional computer graphics. In *Archaeology and computers symposium, Saga Museum, 28–29 May 1991.* Ibaraki: Ibaraki University.

Rahtz, S. & J.Richards (eds) 1989. *Computer Applications and Quantitative Methods in Archaeology 1989.* Oxford: British Archaeological Reports.

Rains, M. 1989. Computerising the plan database: the Aegis system. *Archaeological Computing Newsletter* **21,** 1–2.

Reilly, P. 1988a. *Computer Analysis of an Archaeological Landscape: medieval land division on the Isle of Man.* Oxford: British Archaeological Reports.

Reilly, P. 1988b. *Data Visualisation: recent advances in the application of graphic systems to archaeology.* Winchester: IBM UK Scientific Centre.

Reilly, P. & S.Rahtz (eds) 1992. *Archaeology and the Information Age.* London: Routledge.

Reilly, P. & S.Shennan 1989. Applying solid modelling and animated 3-dimensional graphics to archaeological problems. In *Computer applications and quantitative methods in archaeology 1989,* Rahtz, S. & J.Richards (eds), 157–66. Oxford: British Archaeological Reports.

Scollar, I. 1969. A program for the simulation of magnetic anomalies of archaeological origin in a computer. *Prospezioni Archeologische* **4,** 59–83.

Smith, I. 1985. Romans make a high-tech comeback: Sid and Dora's bath show pulls in the crowd. *Computing* **June,** 7–8.

(Soprintendenza Archeologia Per le Provincie di Sassari e Nuoro) 1989. *S.I.P.I.A.—Progetto S.I. T.A.G. (Sistema Informative Territoriale Archeologico Gallura) Archaeologia del Territorio. Territorio dell' Archeologia: immagini di un' esperienza di catalogazione informatica dei beni culturali della Gallura.* Chiarella-Sassari: Tempio Pausania.

Stančič, Z. 1989. Computervision: producing intra-site plans. *Archaeological Computing Newsletter* **20,** 1–10.

Stewart, W.K. 1991. Multisensor visualization for underwater archaeology. *IEEE Computer Graphics and Applications* **March,** 13–18.

Todd, S.J.P. 1990. *Winchester Colour and Texture Facilities: WINCAT.* Winchester IBM UK Scientific Centre.

Trigger, B.G. 1989. *A History of Archaeological Thought.* Cambridge: Cambridge University Press.

Tyrell, J., F.Yazdy, M.Riley & N.Winterbottom 1990. CARVUPP: Computer-Assisted Radiological Visualisation Using Parallel Processor. Paper presented at Transputer Applications '90, University of Southampton.

Weiss, A. 1989. The compass system. In *Computer applications and quantitative methods in archaeology 1989,* Rahtz, S. & J.Richards (eds), 295–318. Oxford: British Archaeological Reports.

Wood, J. & G.Chapman (with contributions by K.Delooze & M.Trueman) 1992. Three-dimensional computer visualization of historic buildings—with particular reference to reconstruction modelling. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 11. London: Routledge.

Woodwark, J.R. 1986. *Computing Shape.* London: Butterworth.

Woodwark, J.R. 1991. Reconstructing history with computer graphics. *IEEE Computer Graphics and Applications* **January,** 18–20.

Zingarelli, D. 1989. Baldassare Conticallo: cronaca di un progetto. *Rivista IBM,* **XXV(1),** 30–39.

Analysis

13

The simulation and recovery of archaeologically meaningful levels

TODD KOETJE

Introduction

Intrasite spatial analysis in archaeology has traditionally been focused on those sites or levels within them that are most amenable to analysis as if they occupied only a horizontal plane (Whallon 1973; Whallon 1974; Whallon 1984; Koetje 1987a; Koetje 1987b; Simek 1984b; Koetje 1989; see also numerous examples in Hietala 1984). Once a site or level has been chosen, vertical patterning within a single level has most frequently been ignored or defined away.

To be interpreted properly, spatial patterning recovered from the vast preponderance of the archaeological record, must logically contain a consideration of object positioning in three-dimensional space. There are three basic factors that make consideration of three-dimensional patterning important. First, sites, even the most 'pristine', are properly conceptualized as deposits that occur on a pre-existing topography. Even the most appealing living floors contain some vertical relief. Second, the deposition of artefact is usually accompanied by the natural and/or artificial deposition of sediment; sites are basically accretionary phenomena. Third, post depositional processes, both natural and artificial, move artefact and sediment in both the horizontal and vertical directions. These general factors combine to create three-dimensional deposits that we recognize as sites. Many sites, of course, will contain complex vertical palimpsests of artefact from a variety of distinct occupational and depositional episodes.

Two of the most common approaches to choosing a site or level for spatial analysis, involve either accepting large-scale geologically defined sediment units, or seeking to break up thicker layers 'by eye', often through the use of scatterplots or vertical projections. Both approaches have serious weaknesses. Grossly defined geological units can span very long time periods, and may include almost any number of distinct depositional or occupation events. Breaking up the units into levels by eye, however, is also fraught with peril. Perhaps chief among its many problems, is the lack of replicability. It is also not uncommon for changes to be very gradual, or otherwise effectively invisible, except on a fairly gross scale. In addition, defining levels in this manner can create substantial practical problems. Efficient ways to go from lines on a vertical projection that group objects into layers, to marking those objects in a database or catalogue which belong to the newly created layers are poorly developed. Defining layers for analysis in this manner can easily lead to mixtures of discretely deposited materials.

An ideal alternative to these methods for choosing analytical layers would focus on three important considerations:

1. replicability.
2. reliability.
3. ease of presentation and analysis.

These factors suggest that quantitative techniques such as cluster and discriminant analysis could be usefully employed in efforts to create appropriate analytical units. This chapter focuses on the roles that cluster analysis can play in the delimitation and study of spatial patterning in three dimensions.

Although a number of clustering techniques have been utilized in archaeological spatial analysis, the K-means technique has perhaps been the most frequently used (eg. Doran & Hodson 1975; Kintigh & Ammerman 1982; Simek & Larick 1983; Simek 1984b; Bricker 1987; Koetje 1987a; Koetje 1987b; Simek 1987; Koetje 1989). This technique is relatively well understood, well behaved, and accessible. One manner in which cluster analysis might be used to help define or create appropriate layers for spatial and other kinds of archaeological analysis, would involve using it to produce clusters which could be further grouped into distinct analytical levels. In order to demonstrate that this technique can reliably recover or reproduce previously defined layers in three-dimensional space (ie. recover a known structure from a given distribution of points in space), a series of simulated archaeological layers were analyzed using a variety of experimental parameters.

This chapter has two main goals. First, to examine the abilities of the K-means technique in separating or recovering layers where two or more distinct layers can be defined. This will show the basic concordance between K-means results and the ‘reality’ of artefact distributions in three-dimensional space. The second goal is to produce estimates of the error in discriminating between layers under various experimental parameters.

K-means is a non-hierarchical, divisive clustering technique. Extensive discussions of its general characteristics and use in archaeological spatial analysis may be found in a variety of sources (Doran & Hodson 1975, pp. 180–84, 235–37; Dubes & Jain 1980; Kintigh & Ammerman 1982; Simek & Larick 1983; Simek 1984a, p. 28; Koetje 1987b, p. 40; Simek 1987) and so need not be presented in detail here. Briefly, the technique works by searching for cluster formations that minimize the global Sum of the Squared Error (SSE), or the squared distance between a cluster’s centre and its constituent members. The result is a set of so called ‘optimal’ cluster solutions, where the SSE is at a minimum as determined by subsequent changes in the assignment of objects to clusters. For example, the 3 and 12 cluster solutions might be found to be ‘optimal’ in a particular analysis. Each of these would group all objects in a distribution into either 3 or 12 clusters. The clusters in a solution are derived independently from the clusters in any other solution. When used to cluster objects by their spatial coordinates, the K-means technique provides the basis for what has been called a locational analysis (Koetje 1989).

As with all clustering techniques that attempt to minimize a distance based error measure, the K-means technique tends to find clusters of equal variance in all dimensions, and is often referred to as an ‘equal variance method’ (Everitt 1980; Spath 1985). In three dimensions an ideal equal variance cluster would be spherical, a shape that does not appeal to archaeologists as a model for deposits. This problem can be overcome through the use of standard mathematical transformations of the spatial relationships under consideration. Fortunately, sphericity in itself is not so much of a problem as is thickness in the vertical dimension. Most archaeologists would be willing to accept a lozenge or highly stretched ovate shape as a general, idealized model of an archaeological level.

In general terms then, a problem with the use of the K-means technique involves finding meaningful clusters in a space of arbitrary dimensional variation while avoiding the problems created by equal variance assumptions. One solution is to utilize some transformation that creates an analytic space with equal variance in all dimensions, while maintaining the relative distance relationships among the objects characteristic of their original distribution. This ‘equal variance space’ (S_1) is an analytic construct, and thus quite distinct from the true or original space (S_0) that the objects inhabit.

The z-score (or standard score) transformation is the most generally useful in producing such an equal variance analytical space. However, dividing the coordinates along each axis by their standard deviation produces a similar result.

After using either of these transformations, the standard deviations of all coordinates in the new, equal variance space are equal to 1.0, and the relative distance relationships within the original data are preserved. In such a situation all dimensions contribute equally to cluster formation. At this stage in the process any other type of transformation is equally acceptable so long as the objects retain their relative distance relationships and the variance of all dimensions is equalized. Metric equivalency between the true space and transformed space need not be maintained (Everitt 1980).

Given transformation into an equal variance space, and some substantial object density, K-means will generally be successful in finding roughly spherical clusters. There are two important characteristics to notice, however. First, any clusters are only as spherical as the data allow; the actual point cloud of any given cluster may take almost any shape because all points are included at a given solution. Second, the sphericity is only readily apparent in the equal variance space (S_1): when cluster contents are mapped back into their true space (S_0), much of this sphericity disappears.

Given equal variance in the two horizontal dimensions and 100 times less variability in the vertical dimension (for example a volume with dimensions 10(×1000×10), spherical clusters formed in an equal variance space (S_1) will appear lozenge or ovate shaped when mapped back into the true space (S_0). The SSE contribution of the Z axis (depth) can effectively be scaled to equal that of the X and Y axes for the clustering procedure. These characteristics suggest that the potential of K-means clustering for finding archaeologically interesting clusters in three dimensions is high. In an effort to discover the best parameters for such an effort, a series of test runs were performed on simulated assemblages.

Experimental parameters for clustering simulated assemblages

While there have been a large number of similar experiments designed to demonstrate various properties of clustering techniques (Gower 1967; Sneath & Sokal 1973; Kuiper & Fischer 1975; Dubes & Jain 1976; Mezzich 1978; Dubes & Jain 1979; Bayne, Beauchamp, Begovich & Kane 1980; Dubes & Jain 1980; Everitt 1980; Milligan 1980; Milligan 1981; Raghaven & Ip 1982; Spath 1985; Willet 1987), none have been published that concentrate on clustering in archaeological spatial analysis. The experiments described below were designed to explore the effects of four general parameters on the recovery of objects grouped into two or more discrete layers. Three major assumptions are involved in this effort. First, that the ability to recover the layers, or at least to discriminate between them has some meaning, and is the general goal of the process. Second, that the layers to be recovered are horizontally oriented, and third, that individual clusters or groups within a given layer are also horizontal in orientation. Thus while sloped distributions are not directly addressed in the experiments below, any distribution with a reasonably consistent slope, or a characteristic three-dimensional surface that can be described adequately, can easily be transformed into an effectively horizontal analytic space through a variety of techniques. The use of regression or trend surface analysis residuals as offsets from a horizontal plane is among the more obvious methods of creating this horizontal analytic space.

In these experiments, layers were intended to simulate simple archaeological levels. Each layer consists of a distinctly labeled, unique set of three objects distributed in a 1000×1000×10 unit volume. In each experiment, the parameters were held constant over at least three independently created sets of layers. Whole unit X, Y and Z axis coordinates for each layer were generated independently of all others.

The objects in a given layer set were formed into from 2 to 60 clusters using the K-means technique. Error in discriminating between the layers was measured using the number of misassigned objects (MAO). Any object that appeared in a cluster whose membership was predominantly from another layer was counted as a misassigned object. In the case of clusters that had equal numbers of objects from more than one layer, the number of misassigned objects was the cluster total divided by the number of layers represented.

Although a number of experimental parameters were investigated, only the two most critical will be discussed here. The first experimental parameter was the distance separating the layers along the Z axis (vertical). This was independently and systematically varied from 50 to 1 unit(s) for each set of layers. The second experimental parameter investigated involved the multiplication of the Z axis coordinates by a constant, called the expansion factor.

Various expansion factors were applied to the Z axis coordinates in both the transformed and untransformed spaces.

Experimental results

The experiments can be conveniently formed into four groups based on the interpretive goals, and patterning involved in each. Each group includes a number of series, consisting of all those individual runs sharing a set of experimental parameters.

Table 13.1 Experimental results for groups 1–3

Variable	Group		
	1	2	3
1	None Z-Score S.Deviation	Global: Z-Score S.Deviation	Within Layer:
2	1–50 units	10–50 units	1–10 units
3	yes	yes	yes
4	various	none	none
Result:	very poor	excellent to perfect	excellent to perfect

Variable key:
1. transformations
2. layer separation
3. clustered/unclustered layers used
4. Z axis distortion

Group 1

Group 1 consisted of exploratory runs, in an attempt to establish some basic expectations for the procedures. No transformations were used, but various expansion factors were applied to the Z axis coordinates. In general none performed even moderately well with layer separation factors as high as 50 units. In no case was there perfect discrimination between layers (ie. no misassigned objects). The number of misassigned objects often amounted to between 30 and 50% of the total number of objects in both layers at any given cluster solution (see Table 13.1).

Group 2

Group 2 focused on exploring the effects of the Z-score transformations on discrimination of two layers with separation factors ranging between 10 and 50 units. The same results were achieved regardless of the type of transformation used. Discrimination was invariably perfect after the 4 cluster solution, and often after the 2 cluster solution (see Table 13.1).

Table 13.2 Summary of Group 4 experimental series.

Series	# Clustered Layers	Vertical separation (in units)	Mean MAO (csl 30, dz5*)
1	0	2	0
2	0	1	6
3	1	2	0
4	1	1	2
5	2	2	0
6	2	1	4
7	2	2	0
8	2	1	1

* Mean MAO for a given series at the 30 cluster solution, with a Z-axis expansion factor of 5

Group 3

Group 3 focused on the effect of using as much *a priori* information as possible. Each layer was transformed by its *within layer* z-score. The experiments in this group, like those in group 1, were aimed at providing some baseline results. The use of within level z-score transformations would be expected to produce perfect or near perfect results unless the two layers overlapped each other thoroughly. Layer separation was varied from 10 to 1 unit(s). All series showed perfect discrimination (see Table 13.1).

Group 4

The experiments in Group 4 focused on attempting to increase the ability to discriminate between two layers with vertical separation ranging from 5 to 1 unit(s), while using only *global* transformations. All runs with 5 units of vertical separation between the layers achieved perfect discrimination at solutions with two or more clusters (see Tables 13.2–13.3).

Series #6 is the most interesting in group 4. In this series the separation between layers was reduced to 1 unit. In general the ability to discriminate between the layers is reduced in this situation as compared to the other runs in group 4. Although the highest MAO (101) is some 16% of the total number of objects, a much more typical value would be below 10 for any expansion factor of 2 and above. Clearly even in this ‘worst case’ the method does a reasonable job detecting the two layers under these assumptions (Table 13.4).

These data make it possible to generate error estimates for cluster analyses made under particular assumptions. General error rates can be extrapolated from Tables 13.1–13.4, which represent selected results of a more extensive set of experiments.

Table 13.3 Group 4. Average number of misaligned objects (MAO) by cluster solution and Z axis expansion factor.

Number of Misassigned Objects

Cluster Solution

Z	60	55	50	45	40	35	30	25	20	15	10	5
Seriees #1, 2 unit.s of sepparation, no cclusters												
1	10	12	11	11	14	15	15	12	11	5	8	77
2	0	0	0	1	1	0	0	0	0	8	37	1
5	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
Series #2, 1 unit of separation, no clusters												
1	27	29	29	30	35	33	37	45	37	36	12	94
2	14	15	16	18	14	18	18	16	16	14	45	36
5	3	0	2	2	2	2	6	15	15	21	3	0
10	0	0	0	0	1	1	2	2	0	0	0	48
12	0	0	0	0	0	0	0	0	0	0	0	0

Series #3, 2 units of separation, 1 clustered layer

1	6	5	7	9	10	12	12	13	3	3	0	14
2	1	2	2	2	4	2	2	7	8	3	6	8
5	0	0	0	0	0	0	0	0	0	0	0	1
10	0	0	0	0	0	0	0	0	0	0	0	0

Series #4, 1 unit of separation, 1 clustered layer

1	14	14	14	13	21	18	13	16	16	13	17	17
2	8	10	8	9	9	9	8	10	11	19	8	25
5	2	2	2	2	2	2	2	2	2	4	1	2
10	0	0	0	0	0	0	6	6	0	0	0	26
12	0	0	0	0	0	0	0	0	0	0	0	0

Series #5, 2 units of separation, 2 clustered layers

1	1	4	4	9	13	13	12	16	16	11	8	68
2	0	0	0	0	0	0	1	1	5	12	7	0
5	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0

Table 13.4 Group 4 series #6. Number of misassigned objects (MAO) by cluster solution and Z axis expansion factor at 1 unit of vertical separation and 2 similarly clustered layers, by layer set.

Number of Misassigned Objects

Cluster Solution

Z	60	55	50	45	40	35	30	25	20	15	10	5
Layer Set A												
1	23	23	23	28	29	29	29	26	32	26	26	49
2	14	15	20	20	13	10	12	18	9	17	18	23
4	10	10	10	20	4	4	6	5	7	6	4	0
8	2	2	2	2	2	2	0	0	3	4	28	51
10	0	0	0	0	0	0	0	0	0	6	0	29
12	0	0	0	0	0	0	0	1	2	7	7	27
16	0	0	0	0	0	0	0	0	0	0	0	27
18	0	0	4	4	12	20	21	20	24	25	23	21
20	0	0	0	0	0	0	0	0	0	0	0	27
25	0	0	0	0	0	0	0	0	0	0	0	42
Layer Set B												
1	25	31	30	36	31	31	40	49	39	45	34	68
2	16	17	20	21	22	20	23	18	9	17	33	56
4	6	6	5	5	5	6	7	9	10	7	8	60
6	3	3	4	4	4	4	4	7	4	34	34	34
8	0	7	7	7	7	7	7	4	4	4	31	29
10	0	0	0	0	0	0	0	0	0	6	0	29
12	0	0	0	0	4	4	4	4	4	4	0	29
15	0	0	0	0	21	21	29	29	29	25	29	29
18	0	0	4	4	29	29	29	29	29	25	29	34
20	0	0	0	0	29	29	29	29	29	25	29	34
25	0	0	0	0	0	0	0	0	0	0	0	60

Layer Set C												
1	41	41	43	43	43	43	36	41	38	38	101	31
2	16	18	20	20	22	20	23	18	9	17	33	82
4	6	6	7	5	2	4	5	5	8	4	7	32
6	2	2	2	4	4	2	2	2	3	0	13	68
8	0	0	0	0	0	0	2	2	5	5	12	68
10	24	26	26	26	26	26	26	28	28	31	31	68
12	0	0	0	0	0	0	0	0	0	2	0	68
14	0	0	0	0	0	0	0	0	0	0	0	31
18	0	0	0	0	0	0	0	0	0	0	0	68
20	0	0	0	0	0	0	0	0	0	5	29	31
25	0	0	0	0	0	0	0	0	0	0	31	31

Discussion

Several points seem clear as a result of the clustering experiments. First on both theoretical and empirical grounds, the use of *within layer* z-score transformations is the best solution to the problem of discriminating between two or more layers regardless of their vertical separation.

This relies however, on *a priori* knowledge about the layers that is quite specific: a reliable layer assignment must be available for each object. Clearly this begs the question. In addition, it would be equivalent to separate analyses done within each layer. There is, however, value in a more holistic approach even given such precise knowledge. Several basic relationships may be profitably explored, leading to:

1. exposure of those areas or volumes in a site that might have questionable layer attributions, may be mixed or are distributionally anomalous in some other sense.
2. allowing subsequent analyses to crosscut previously defined layer boundaries in these areas.

Minimally then, this approach might be used to create a replicable verification of more traditionally defined layer boundaries, especially when such boundaries are not visually obvious. In essence such a use would examine within versus between layer variation, with outlying groups targeted for closer examination. These outlying groups of objects might correspond to:

1. distinct depositional levels, or depositional/disturbance episodes within a particular sediment layer.
2. differential post-depositional behavior of material classes within a single sediment layer.
3. objects that have been mislabeled or mismeasured.

The second main point to be gleaned from the experimental results is that using global z-score transformations and Z axis expansion factors between 2 and 30 can be nearly as successful as within layer transformations in discriminating between layers regardless of vertical separation or other layer characteristics. Expected error rates are somewhat higher than those produced when using within layer transformations, but are certainly within reasonably acceptable limits in almost all cases. This is especially important considering the limited knowledge about a distribution required to use global transformations. The experimental results demonstrate that for a variety of circumstances, Z-axis expansion factors of five or greater almost always lead to error in object assignments of approximately 5% (of the total number of objects) or less. Only rarely was this error in fact greater than 1%. While this rate is not insignificant, it is certainly acceptable, especially given that there are no published error estimates for the more traditional methods. As a completely subjective, and probably biased, estimate of error in making layer or level assignments through traditional means I would suggest a range of from 1 to 10%, making it roughly comparable to that found for quantitative methods. We may note that when the Z-axis separation is one unit, visual separation of the experimental layers is effectively impossible, making the error rate 100%.

The simulation experiments presented here indicate generally that these methods are appropriate for spatial analysis when the target layers to be recovered (or perhaps created) can be assumed to be, or defined as, essentially horizontal. In most of these situations, layers may be recovered with an error rate of less than 5% of the total number of objects. These methods may be used to discover and characterize, in a quantified, replicable manner, the basic spatial patterning at a site in three dimensions. Methods of this general type have the potential to vastly increase the number of sites that can be included in our study of intrasite spatial patterning. This type of analytic method has the potential to increase the resolution with which we view intrasite spatial patterning.

References

Bayne, C.K., J.J.Beauchamp, C.L.Begovich & V.E.Kane 1980. Monte Carlo comparisons of selected clustering procedures. *Pattern Recognition* **12** 51–62.

Bricker, H.M. 1987. Artificial habitation structures in the Chattelperronian: Evidence from Les Tamborets (Haute-Garrone, France). Paper presented at the Cambridge Conference on Human Origins.

Doran, J. & F.R.Hodson 1975. *Mathematics and Computers in Archaeology*. Edinburgh: Edinburgh University Press.

Dubes, R. & A.K.Jain 1976. Clustering techniques: the user's dilemma. *Pattern Recognition,* **8,** 247–260.

Dubes, R. & A.K.Jain 1979. Validity studies in clustering methodologies. *Pattern Recognition* **11,** 235–254.

Dubes, R. & A.K.Jain 1980. Clustering methodologies in exploratory data analysis. *Advances in Computer Archaeology* **19,** 69–181.

Everitt, B.S. 1980. *Cluster analysis.* London: Heinemann.

Gower, J.G. 1967. A comparison of some methods of cluster analysis. Biometrics, **23,** 623–637.

Hietala, H. (ed.) 1984. *Intrasite spatial analysis in archaeology.* Cambridge: Cambridge University Press.

Kintigh, K.W. & A.J.Ammerman 1982. Heuristic approaches to spatial analysis in archaeology. *American Antiquity* **47,** 31–63.

Koetje, T.A. 1987a. Pavement features from the European upper paleolithic. Paper presented at the 52nd Society for American Archaeology, Toronto.

Koetje, T.A. 1987b. *Spatial patterns in Magdalenian open air sites from the Isle Valley, southwestern France.* Oxford: British Archaeological Reports.

Koetje, T.A. 1989. A preliminary look at three dimensional spatial patterning: the case of le Flageolet II (Dordogne, France). Paper presented at the 54th Society for American Archaeology, Atlanta.

Kuiper, F.K. & L.Fischer 1975. A Monte Carlo comparison of six clustering procedures. *Biometrics* **31,** 777–783.

Mezzich, J.E. 1978. Evaluating clustering methods for psychiatric diagnosis. *Biological Psychiatry* **13,** 265–281.

Milligan, G.W. 1980. An examination of six types of error perturbation on 15 clustering algorithms. *Psychometrics* **45**, 325–342.

Milligan, G.W. 1981. A review of Monte Carlo tests of cluster analysis. *Multivariate Behavioral Research* **16,** 379–237.

Raghaven, V.V. & M.Y.L.Ip 1982. Techniques for measuring the stability of clustering: a comparative study. *Lecture notes on Computer Science* **146,** 209–237.

Simek, J.F. 1984a. Integrating pattern and context in spatial archaeology. *Journal of Archaeological Sciences* **11,** 405–420.

Simek, J.F. 1984b. *A K-means approach to the analysis of spatial structure in upper paleolithic habitation sites: Le Flagelot and Pincevent 36.* Oxford: British Archaeological Reports.

Simek, J.F. 1987. Spatial order and behavioural change in the French palaeolithic. *Antiquity* **61,** 25–40.

Simek, J.F. & R.Larick 1983. The recognition of multiple spatial patterning: a case study from the French upper paleolithic. *Journal of Archaeological Science* **10,** 165–180.

Sneath, P.H.A. & R.R.Sokal 1973. *Numerical taxonomy*. San Francisco: Freeman.

Spath, H. 1985. *Cluster analysis and dissection.* New York: Springer Verlag.

Whallon, R. 1973. The spatial analysis of occupation floors I: the application of dimensional analysis of variance. *American Antiquity* **38,** 266–278.

Whallon, R. 1974. Spatial analysis of occupation floors. II: The application of nearest neighbour analysis. *American Antiquity* **39(1),** 16–34.

Whallon, R. 1984. Unconstrained clustering for the analysis of spatial distributions in archaeology. In *Intrasite spatial analysis in archaeology,* Hietala, H. (ed.), 242–277. Cambridge: Cambridge University Press.

Willet, P. 1987. *Similarity and clustering in chemical information systems.* Letchworth: Research Studies Press.

14

Current information technology applications to archaeological data from Lower Nubia

PAUL SINCLAIR & LANA TROY

Introduction

The massive rescue operation of the archaeological sites of Lower Nubia (Fig. 14.1), undertaken in connection with the construction of the High Dam at Aswan in the 1960s, has produced an immense amount of new archaeological data, covering every period from palaeolithic times to the medieval Christian era (Säve-Söderbergh 1987). The Scandinavian Joint Expedition (SJE) to Sudanese Nubia (1960–1964), jointly staffed and financed by Denmark, Norway, Finland and Sweden, and led by Professor Torgny Säve-Söderbergh, was responsible for an area stretching along the east bank of the Nile from the Egyptian border in the north to Gamai in the south, as well as the west bank tomb of Amenemhet, an area of c. 150^2 km.

Some 490 sites, containing c. 4200 tombs representing periods from the palaeolithic to middle ages, were found in the SJE concession area. The results of this expedition have been published in Volumes 1–9 in the SJE series. The last sites to be published have been those belonging to the Middle Nubian and New Kingdom periods (Fig. 14.2). The publication of these sites (Vols 4:1–2 and 5:2–3) has been delayed to the end of the 1980s and the beginning of the 1990s, a delay which has put this material within the grasp of the ‘age of the microcomputer’.

The Middle Nubian and New Kingdom periods (c. 2200–1300 BC) are of special interest to the study of the history of the region (eg. Säve-Söderbergh 1941; Adams 1977.) ‘Middle Nubian’ (c. 2200–1500 BC) is a collective term applied to the indigenous cultural groups of Lower Nubia, more specifically the area between the first and second cataract. Although the Egyptian presence in Lower Nubia is significant during this time, consisting of a chain of fortresses overseeing the exploitation of natural resources during the Middle Kingdom (ca 2000–1700 BC), the finds from the Middle Nubian sites, primarily cemeteries, largely reflect the burial customs of the indigenous population, documented in previous studies as the C-Group and Pangrave peoples (Bietak 1966; Bietak 1968; Bietak 1987).

Following the end of the Middle Kingdom, during the Second Intermediate Period (c. 1700–1550 BC), the Upper Nubian kingdom of Kush, with its centre at the third cataract site Kerma (Bonnet 1978; Bonnet 1980; Bonnet 1982; Bonnet 1984; Bonnet 1986; Gratien 1978), had control over Lower Nubia. The features which distinguish the Kerma cultural group are only represented by isolated graves and a few artefact types in Lower Nubia.

With the defeat of the Kushites by the Egyptians in c. 1550 BC, the archaeological evidence, derived almost exclusively from cemetery sites, indicates a rapid adoption of burial customs which differ significantly from those familiar from the Middle Nubian period. The funerary practices appear as ‘Egyptianized’. The contracted burials in circular pits with rings walls of various types of the indigenous Middle Nubian peoples become extended burials in rectangular shafts and burial chambers. The handmade Nubian ware, often black-topped or with incised decoration, is exchanged for the mass produced Egyptian wheelmade vessels, sometimes painted but more often plain. The elaborately beaded loincloths and leather sandals are abandoned for the funerary trappings of the coffin and funerary mask.

The Lower Nubian cemeteries of the New Kingdom period both parallel the burial customs of their northern neighbors and differ in their lack of objects such as ushebtis, stelae and other items which are dependent on a knowledge of the hieroglyphic script and reflect the theological core of the Egyptian funerary rituals.

The sites from these two periods found within the SJE concession offer unique opportunities for a detailed examination of the quantitive variations represented by cultural groupings in the Middle Nubian material and in a chronological continuum in the New Kingdom data. And it was this material, with its well defined quantification, to which multivariate analysis was applied.

The classical contributions of Flinders Petrie to archaeological seriation, made at the end of the 19th century, have long been recognized as fundamental elements in archaeological methodology. In the intervening years analyses of the contents and geographical layouts of cemeteries in Upper Egypt were carried out by such successors to Petrie as Brunton and Reisner (cf. Kemp’s discussion (Kemp 1976, p. 265 ff.)). In more recent times, with the explosive development of the ‘New Archaeology’ in the 1960s, a range of new techniques for cemetery analysis and seriation, often with a focus on European and American data, have been made available to those working with archaeological data from Egypt and Nubia. These

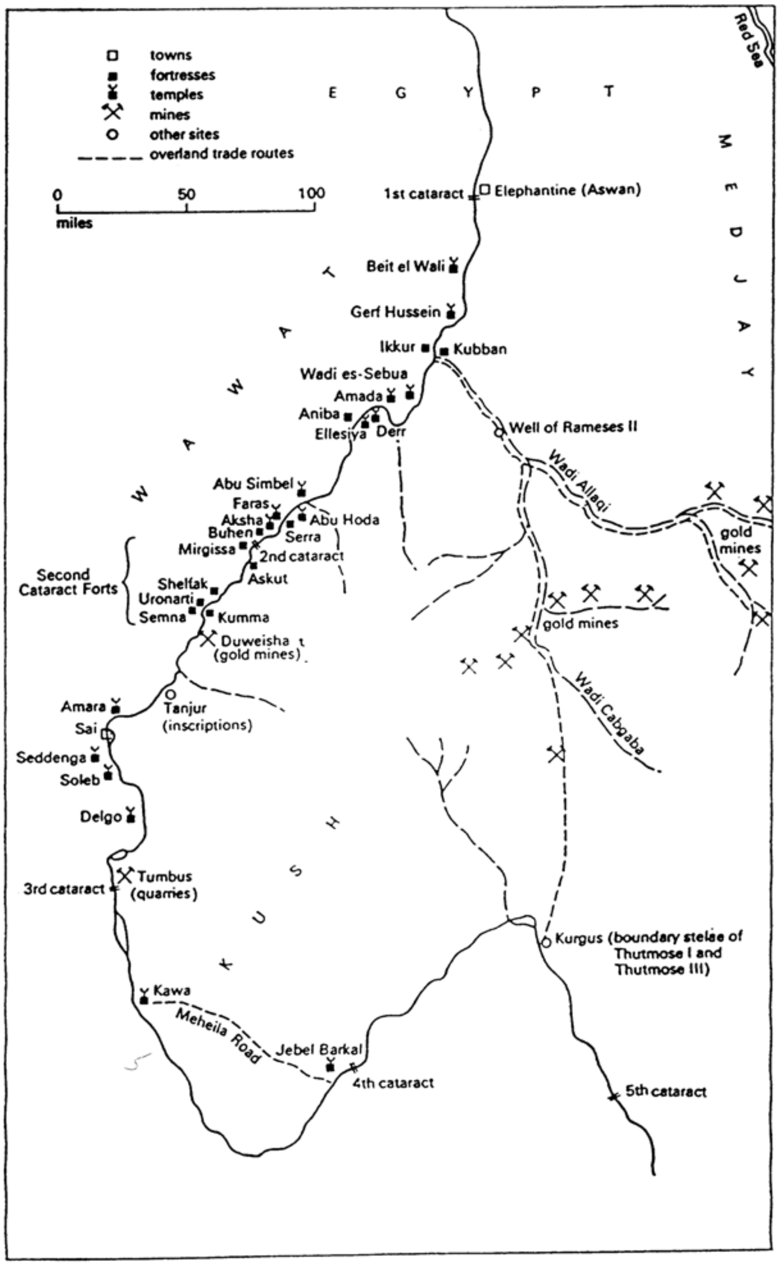


Figure 14.1 Lower Nubian sites.

developments, and especially those of relevance to Egyptologists, were succinctly summarized by Kemp in 1975 who further suggested a number of procedures for tackling the great complexity obtained from cemetery incidence matrices. In addition to the strictly seriational problems found in the analysis of Egyptian cemeteries, use has been made of the geographical aspects of tomb and artefact location within cemeteries. This approach is exemplified by the work of Manfred Bietak (eg. Bietak 1968) which documents cemetery growth in Nubia. In 1975 Kemp was however forced to conclude that even a moderate size matrix required a number of permutation operations 'beyond the reach of the most advanced high-speed computer which can be envisaged' (Kemp 1976). In the 70s the goal of obtaining a perfectly sorted Petrie Matrix seemed at best formidable. Since then multivariate analysis has had a limited application to Egyptian and Nubian archaeological material.

Accompanying this trend is the increasing difficulty in communicating results in a manner which is meaningful to other specialists. The thought remains that the broader significance of the different specialist contributions is to be seen in terms of correlation with other categories of evidence. In assisting with the compilation of results from SJE the need for an integrative approach to analyzing complex data sets has become increasingly apparent. Developments in the last decade in the archaeological application of multivariate analysis in Scandinavia and particularly that of correspondence analysis (Benzecri 1973; Bolviken, Helskog, Helskog, Holm-Olsen, Solheim, & Bertelsen 1982) have provided a strong stimulus for exploratory

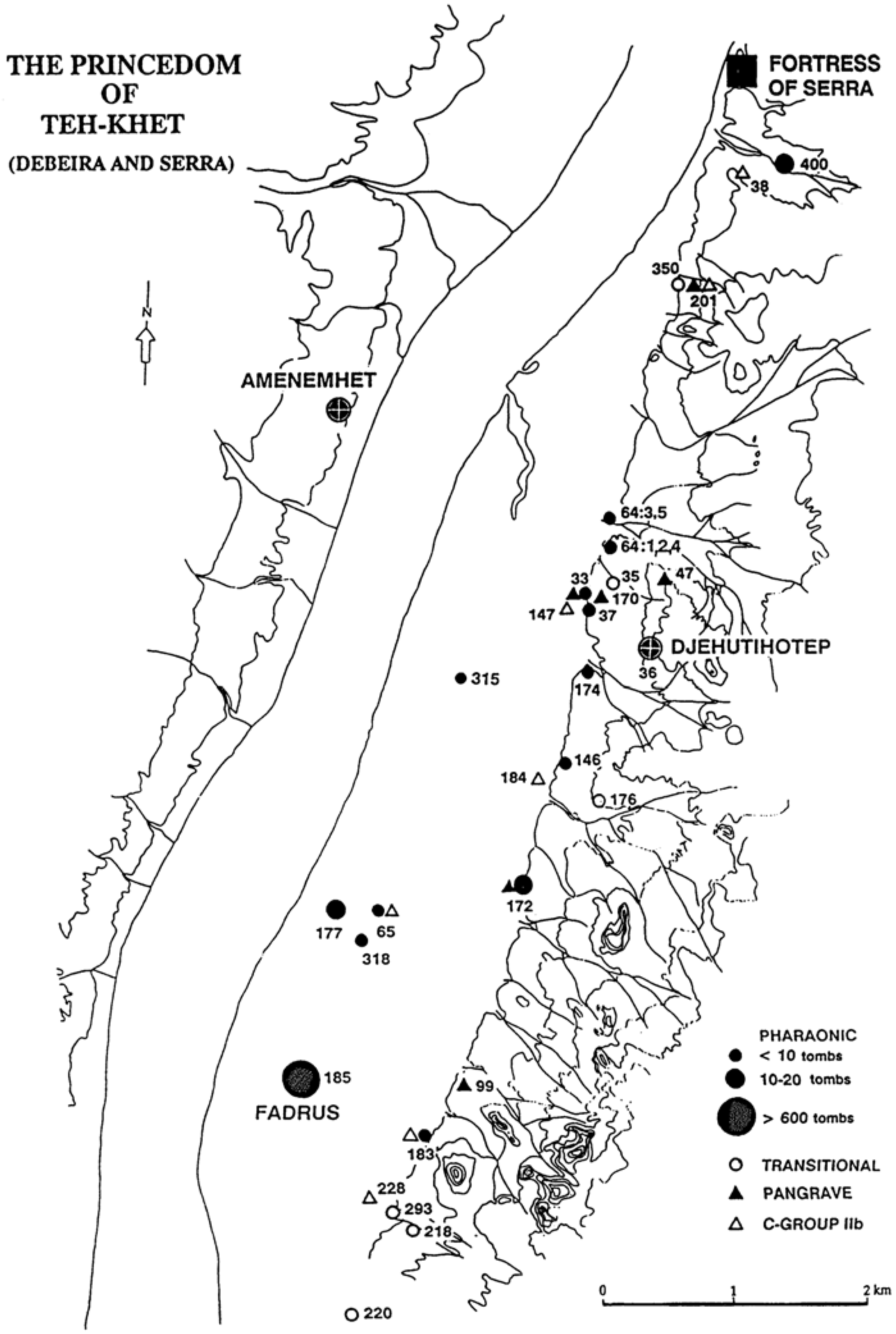


Figure 14.2 Middle Nubian and New Kingdom sites in the SJE concession area.

analyses of Nubian material. The following examples are not exhaustive but are intended to show the potential of the application of multivariate statistical techniques both in integrating previously separate analytical approaches and in providing new hypotheses about complex data sets.

Preparation had been made for the application of computer techniques in the early phases of the work on publications covering Middle Nubian and New Kingdom sites (Sinclair & Troy 1988; Sinclair & Troy 1989).

The Middle Nubian analysis—C-Group, Pangrave and transitional sites

The SJE concession contained a total of sixty-one Middle Nubian cemetery sites. Over 2000 graves were found on these sites. The finds suggested a dating from the Middle Kingdom to the middle of the 18th Dynasty (c. 2000–1450 BC). As the preservation of these sites was poor, with a large percentage of the graves plundered and the original attribution of finds uncertain, the study treated the site, rather than the individual units, as the object of analysis.

Twenty-four of the sixty-one sites were chosen for inclusion in the analysis on the basis of adequate attribute representation. These sites contained c. 75% of the total Middle Nubian grave units and almost all of the surviving artefacts. The twenty-four sites were characterized 'culturally' according to their most dominant features as belonging to the C-Group (eight sites), Pangrave (five sites) or Transitional (five sites, cf. Vol. 4:1) groups. The latter group of sites have been described in terms of the combination of traditional Middle Nubian burial customs (tomb type and body position) and the gradual domination of Egyptian grave goods. These sites are placed chronologically in the 18th dynasty (Säve-Söderbergh in Vol. 4:1, pp. 23 f.) and thus are contemporary with the New Kingdom sites. In addition four sites of mixed elements, representing different combinations of C-Group, Pangrave and Kerma attributes, as well as two sites of ambivalent Middle Nubian character were included in the analysis. The identification of these sites, according to cultural affinity, was made using comparative material and was not a determinative factor in the analysis.

For the abstraction of quantitative data, three main areas were chosen. Tomb construction and burial place, pottery and other (non-ceramic) finds. Each of these areas had further subdivisions. For example, placed under the heading 'Tomb construction and burial places', were the subdivisions, superstructure types, superstructure materials, offering places, shaft types, shaft section, burial place type, body position, gender distribution and head orientation. The heading 'Pottery' took up large groups of handmade pottery types, as well as distinctive wares and decors. 'Other finds' covered the area of types of finds such as beads, scarabs, kohl pots, spacers and ostrich feathers, to name a few, as well as bead material (ostrich egg shell, faience, shell etc.) and find materials such as wood, leather, textile and precious metals. In this way a total of ninety-three attributes for the twenty-four sites were determined as the basis for an incidence matrix. The matrix was built up using the quantity of each attribute per site as it was abstracted from databases covering tomb construction, handmade ceramics and wheelmade ceramics.

The analysis was run with the CANOCO program for correspondence analysis (Ter Braak 1988), on a Toshiba 1200 microcomputer with a maths coprocessor. The goal of this analysis was to present a visual representation of the relationship between the various cemeteries and their attributes. This was undertaken both as a test of the hypothesis of cultural groupings (C-Group, Pangrave and Transitional) and to examine the association of the various attributes with these groupings. The result of this analysis was presented in Vol. 4:1 in a series of figures where the position of the sites and the attributes had been placed on the basis of the first two axes generated by the CANOCO program.

The results were of great interest (Fig. 14.3), clearly showing the validity of the grouping of the sites termed 'Transitional' (Fig. 14.3 site numbers are underlined as preceded by T, lower righthand corner), as well as the distribution of Pangrave sites (preceded by P, upper lefthand corner) and the clustering of C-Group sites (middle of figure). The positioning of various attributes both illustrates and suggests their cultural placement. In the figure the abundance of each attribute varies linearly from the origin of the axes (lowest values) to its position on the figure (highest value). The C-Group attribute, hair clasps (attribute no. 75), is positioned in the C-Group site cluster, as is the attribute representing ostrich feathers (no. 84), recognizable from the distribution of the concession finds as characteristic of the C-Group. The few kohl pots (no. 78) are associated with Pangrave sites, while the placement well in the centre of the C-Group sites, of Nerita shells (no. 81), often regarded as a distinctively Pangrave feature, draws attention to the wider use of this item within a Middle Nubian context. The interaction of the total combination of attributes and sites is illustrated in the placement of amulets (no 73) in the lower righthand corner, indicating its association with both the Transitional sites and with the outlier C262, which has the largest number of amulets of any site.

The application of multivariate analysis to the Middle Nubian sites presented an overall picture of the interrelationship of the three general groups of sites and the attributes associated with them. As to be expected from an analysis which incorporates so many variables, the results reflect both socio-cultural and temporal variation. The similarity produced by cultural and temporal fashions interacts in the projection of the sites and the variables onto a two-dimensional picture.

The specific interpretation of the relationship between individual archaeological dimensions and mathematical axes remains problematic. Complex trends in past behavior, which involve stylistic and socio-cultural variability and result in a variation of abundance of remains in archaeological context, have often, through necessity, been artificially compressed by archaeologists into analytical concepts of 'time' and 'status'.

The traditional use of seriation in archaeology in which 'time trends' are held to provide the most significant source of variability is a case in point. No matter how justified this assumption often is (especially when a relatively limited set of attributes is analyzed), in the case of the multivariate analysis of the Middle Nubian sites, which saw the inclusion of a much wider range of attributes, a more holistic consideration of cultural similarities and differences seems to have determined the outcome of the analysis.

As we will try to demonstrate below, multivariate analysis, and particularly correspondence analysis, provides an opportunity to examine the interaction of such concepts as 'time' and 'status' as they affect complex data sets.

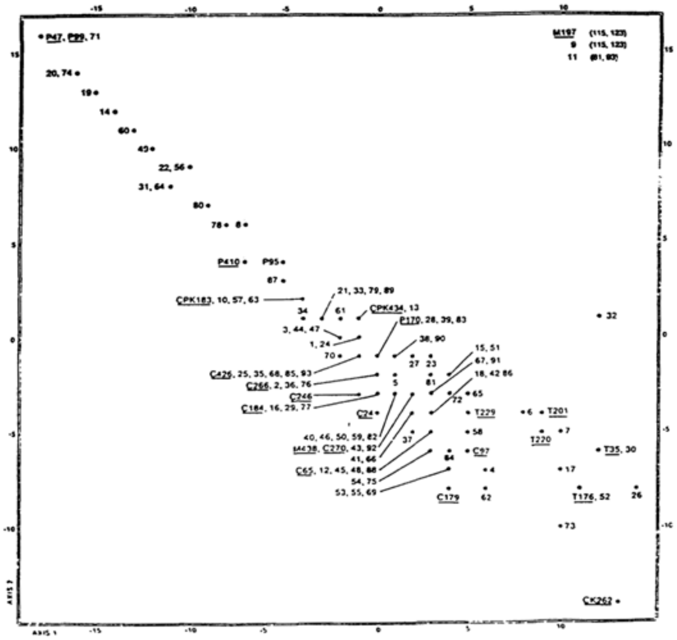


Figure 14.3 Total projection of axes 1 and 2.

The New Kingdom analysis—Fadrus Site 185

While the Middle Nubian analysis concentrated on a number of sites, with the cemetery as the unit of analysis, the starting point of the New Kingdom work was one site, no. 185 (Säve-Söderbergh 1963, pp. 59 ff.), with the individual grave unit serving as the analytic unit.

This site, located on the east bank, c. 650 m from the Nile, on the southern part of the Komangana plain, was situated on a low gravel hillock in the village of Hillet Fadrus. It covered a c. 1200 m^2 area. The centre of the site was covered by the modern houses which occupied c. 16% of its area. The location of the site suggested that it has been associated with a rich farming area, perhaps of importance in supplying the Egyptian military presence. It was, during the New Kingdom, a part of the Princedom of Teh-khet, and ruled by the local Egyptianized prince family, represented within this concession by the tombs of the brothers Djehuty-hotep and Amenemhet.

The cemetery consisted of 715 units, including two low mudbrick surface structures of doubtful origin (Fig. 14.4). c. 91% of the units were undisturbed and contained at least one burial. Unfortunately the skeletal material proved, on the whole, to be too fragile for removal and further examination so that length has been the primary form of documentation of skeletal remains.

The unit types was distinctive. The simple tomb types consisted of simple rectangular shafts (274), side niches commonly with mudbrick sealing (127), end niches also with mudbrick sealing (233), and double end niches with sealing (10) and anomalous corner niches (4). Mudbrick constructions were of two types: mudbrick chambers, with flat or pointed roofs, in shafts with openings from the roof (13) and chambers commonly with vaulted roofs, with doors opening to the west placed in spacious shafts with a pit ('ramp') in front of, or in unusual instances, mudbrick ramp leading to, the entrance (44). In addition there were shafts, with one wall half-way bricked up, possibly to represent a sealing. These have been called 'false side niches' (8) and as they were lacking in burials were probably cenotaphs.

The vast majority of the burials are extended, with only four examples of primary adult burials in contracted position. The extended burials are found on the back (326), right side (142), left side (113) and face (47). The hands were placed either at the side or together at the pelvis, with very few examples at the chest or face, positions common with contracted burials. 92.2% have a west and northwest head orientation. Coffin burials were relatively well represented, although the poor preservation of organic material, due to the flooding of site, makes the documentation of this material uncertain. Similarly, other organic funerary accessories such as textile wrappings and matting were not preserved. A small number of the richer burials were equipped with funerary masks, similar to those found on other smaller New Kingdom sites in the concession area.

Approximately 65% of the units contained grave goods. The most common find was pottery, including 1722 complete vessels and only 427 sherds. The pottery is dominated by the common types Flowerpots, Beerbottles, Cups, Plates, Ovoid jars and Carinated vessels (Family types FP, BB, CU, PL, JO, CS and CV). There is, however, a large variety of vessels represented in smaller numbers.

Other find type categories include seals, figurative pendants ('amulets'), spiral and penannular earrings, bracelets, fingerrings, kohl pots and ointment jars, razors, tweezers, ivory slips from cosmetic boxes, metal vessels, axes, daggers,

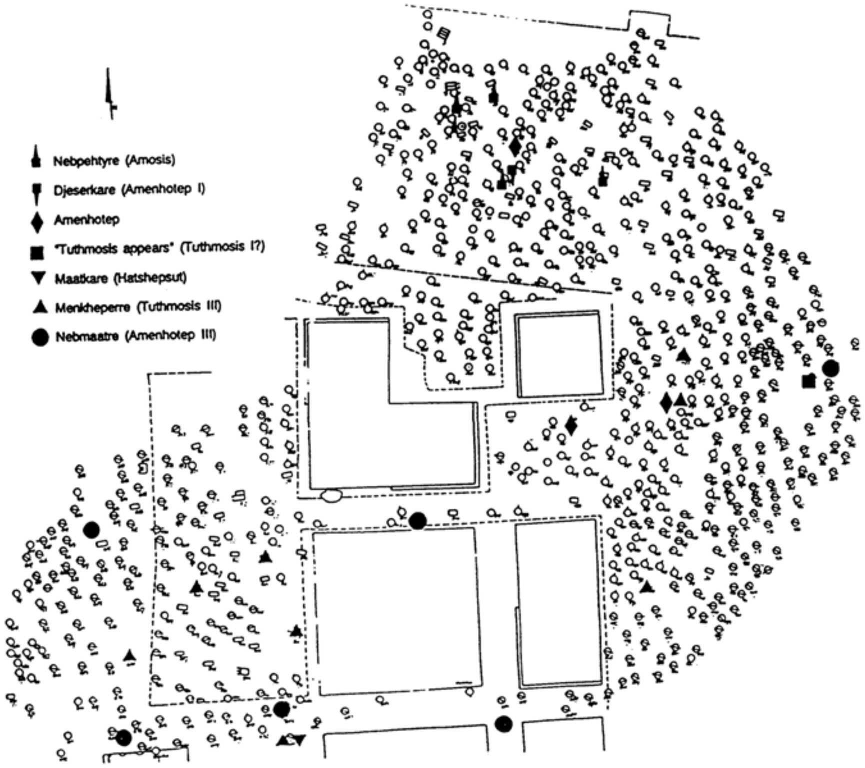


Figure 14.4 Site 185, Fadrus, with the location of seals inscribed with royal names.

arrowheads, fishing hooks and grindstones, as well as bronze fittings of different types. There were also individual finds of objects like a faience bowl, kohl tube, cosmetic saucer, bone awl, an ostrich egg shell vessel and a 'Hyksos' type short sword. And of course beads were represented in a variety of materials and forms.

The finds, on a whole, reflect the common grave goods familiar to both Egypt and Lower Nubia during the 18th dynasty, from sites such as Gurob and Abydos in Egypt and Aniba, Buhen and Semna in Nubia, although notably lacking inscription oriented objects such as stelae, statues and ushebtis.

A small number of seals inscribed with the names of the kings Amosis (c. 1567– 1546 BC), Amenhotep I (c. 1546–1526 BC), Tuthmosis III (ca 1504–1450 BC) and Amenhotep III (1417–1379 BC) provide a tentative guideline for understanding the chronological development of the site, with Amosis and Amenhotep I represented in the northern section, Tuthmosis III in the central section and Amenhotep III in the outlying areas (cf. Fig. 14.4). This general picture tends to be confirmed by the distribution of various ceramic types and by the occurrence of other artefact with association with 'earlier' as opposed to 'later' periods of the dynasty. The combination of ceramic and non-ceramic finds suggests that the cemetery as a whole was in continual use over much of the 18th dynasty.

As with the analysis of the Middle Nubian material, one of the primary aims in quantifying the Fadrus documentation was the pictoral presentation of a very complex set of interrelationships of the material. There were, in addition however, specific problems in relation to the chronological development of the cemetery that it was felt could be addressed within the framework of a multivariate analysis. Primarily there was the question of relating the cemetery growth to the chronology of the 18th dynasty, using the combined attributes of the individual grave units.

An incidence matrix once again provided the starting point. Two analyses were run. The first analysis consisted of 512 units with 171 variables or traits. This analysis is best described as optimal, excluding only those units which either totally lacked finds or only included finds lacking in a significant distribution pattern such as faience beads. An attempt was made to design the matrix, with reference to choice of attributes, in such a way as to include those characteristics which appeared to be of chronological significance. This problem was approached using extensive distribution maps which, with a repeated pattern following the distribution of the seals naming the kings, suggested significant traits.

In the selection of ceramic attributes, the possibilities afforded by the system of typological description devised by Holthoer for the New Kingdom material was exploited, using not only the subfamily types (family variants, such as storage jar subfamilies ST1–4), but also the observations, confirmed by distribution maps, of the varying distribution of wares and decor to further extend the number of chronologically significant attributes. Thus, for example, the cup type CU3 is included in two variations, with ware group IIR (compacted—'early') and ware group IR (uncompacted—'late').

The frequency matrix for this initial analysis of the Fadrus cemetery is 512 by 171. The enormous size of this matrix, far larger than the Middle Nubian example of 24 by 92, exceeded the capacity of the version of the CANOCO program available at the time of this first analysis and consequently the MV-Arch multivariate package written by Wright was used. The matrix was created using the CREATEV program and the correspondence analysis programs BIGCOR1–5 were applied and also the SERIATE and REORDER options. The results of the analysis have provided a matrix reordered both in terms of the sequence of burial units and also of attributes. Even given relatively high eigen values (0.410, 0.405, 0.354, 0.284) the percentage variance accounted for by the first two axes was very small (only 3.7 and 3.6 percent respectively) but this does not seem entirely unreasonable when considering the size of the matrix, its complexity and its wide scatter of low incidence values.

The results of that analysis were examined from a number of different perspectives. A simple seriation was set up using the first axis as the ‘best straight line’. This presented an ordered series of burial units in which proximity of points indicates the relative similarity of the units which they represent. This similarity was interpreted as reflecting the temporal unity of the units. There is (as is common knowledge in seriation) no *a priori* criterion for stating which end of the sequence is earliest in time, even accepting the above mentioned assumption of temporality. In assessing the archaeological value of these results however, it became increasingly obvious that the analysis had indeed discerned temporal trends in this very complex material.

The seriation of the variables revealed a very clear picture (Fig. 14.5), following the pattern suggested by the topographical distribution of the variables. It indicated, for example, the successive replacement of shafts by side niches and later end niches as the most common tomb type, the contrast between the use of the seal as a bead stringed around the neck as an ‘early’ trait and the placement on the hand as a ‘late’ trait reflecting the Egyptianization of the burial customs. Spiral earrings, associated with a non-Egyptian population in Egypt, are placed high on the first axis, while the native version of the Egyptian penannular earring is placed lower on the scale, giving an ‘early’ —‘late’ relationship for the two earring types. But what is even more remarkable is the correct placement of the variables providing obvious temporal guidelines, such as seals inscribed with the names of the kings Amenhotep I (Djeserkare and Amenhotep), Tuthmosis III (Menkheperre) and Amenhotep III (Nebmaatre), confirming that the temporal trend runs from high-early to low-late or right to left. It is important to bear in mind that this type of analysis provides information concerning the variability of different sets of attributes in relation to each other, a perspective which has been previously very difficult for archaeologists to obtain.

An additional touchstone of success for an analysis is the provision of relevant points of reference for archaeologists and new hypotheses for inclusion into the inductive/deductive cycles of description, analysis and interpretation also deserves emphasis here.

The problems addressed by this analysis are to be found on different levels. Although the basic chronology of the cemetery was established by this first step, the question remained as to the extent of the main periods, and whether it is possible to clearly distinguish between them, in what is a long continuum.

Having generated two axes for both the individual units and the attributes, the second part of the analysis continued, accepting as a point of departure the likelihood that the first axis did indeed show a degree of temporal variability subject to the assumptions discussed above. It was then taken as possible to map this temporal variation within the spatial frame of the cemetery. As a first step to achieving this, the locations of the individual burial units were electronically digitized. The resulting list of coordinates was incorporated into files compatible with the Rockware spatial analysis package. The Rockbase module facilitated the building up of a spatially orientated database incorporating the geographical placement of the burial units and the first two axes generated by the correspondence analysis.

The Gridzo module was then used to interpolate regular grid values over the unevenly distributed data and a contour map of the results was produced (Fig. 14.6). The contour map was derived from the positions on the first axis of the correspondence analysis (which as discussed above displayed a chronological relevance). The contours which were produced should then, subject to the above discussed assumptions, be taken to delimit isochronous units. The high values show the early locations in the cemetery and the lower ones, later areas. The contours, drawn over the area covered by modern houses, also projected the hypothetical character of that area in relationship to the values found for neighbouring units. The picture given by the contour map is in general agreement with the dating of different areas of the cemetery generated by distribution maps of the finds. Thus, the contour maps make it possible to see the complete temporal development of the cemetery as a whole.

It is important to bear in mind that the measure of similarity between burial units is dependent upon variation within the list of variables included in the original analysis. Consequently, it is possible to register both similarity between widely separated units and also the progressive replacement of nearby units all within the same diagram. A further possibility exists of using the isochronous contours to reincorporate the burial units initially excluded from the analysis owing to a paucity of finds. But it is also important to underline that the contour map provides a best overall approximation and individual anomalies can be expected to occur.

In order to obtain a more detailed view of the socio-cultural and temporal anatomy of the cemetery, an additional analytical element indicating status ranking was added to the analysis. This was done in an attempt to distinguish high and low status areas distributed both geographically on the digitized map and in relationship to the time zones projected from axis 1 of the correspondence analysis. The status ranking of each unit was determined taking into consideration construction elements of

	CANOCO (474×145)		Wright (512×171)	
FUNERARY ACCESSORIES				
Funerary mask	18	36	2.231	−1.136
Coffin	−12	1	−0.825	−0.458
PERSONAL ADORNMENTS				
Beads				
Precious metal	19	−28	0.897	2.274
Faience	11	−26	0.161	1.961
Ostrich egg shell beads	11	−35	0.253	2.193
Carnelian	10	−28	0.330	2.006
White faience	9	−41	0.148	2.290
Yellow faience	7	−35	0.328	1.554
Red faience	−1	−22	−0.173	0.910
Seals				
Scarabs	17	−27	0.673	2.201
Scaraboid	10	−32	0.367	2.015
Plaquette	10	−35	0.245	1.888
Inscribed with royal (givingterminuspost quem)				
Amenhotep	24	−16	0.979	1.738
Djeserkare	21	−53	0.636	3.641
Menkheperre	−1	−18	−0.204	0.858
Nebmaatre	−25	−8	−1.368	0.288
Position on body				
On neck	20	−32	0.649	2.571
On pelvis	14	−44	0.245	3.154
On hand	6	−15	0.020	1.058
On feet	0	−15	−0.494	0.743
Figurative pendants				
Flies	15	−8	1.101	1.030
Taueret	8	−19	0.530	0.686
Fish	4	−2	0.277	0.438
Hand pendant	4	−6	0.346	−0.091
Poppyheads	2	−7	0.245	0.064
Lily	2	−6	−0.050	0.368
Ducks	1	−12	0.055	0.576
Heart	1	−29	−0.200	1.419
Glass Birds	−15	−35	−1.470	1.820
Other adornments				
Spiral earrings	23	−31	0.520	2.446
Penannular earrings	0	−24	−0.283	1.226
Toilette Equipment Kohl pots				
Tall	35	−25	1.038	2.275
Squat	20	−12	0.812	1.384
Necked	11	0	0.735	−0.041
New Kingdom type	6	2	0.471	−0.050
Ivory inlay for cosmetic boxes	24	4	1.836	0.149
Tweezers	21	5	1.803	0.195
Weapons				
Swords and daggers	30	12	2.260	0.817
Battle axes	25	2	2.516	−0.173

Figure 14.5 Selected list of variables, comparing results from CANOCO and Wright.

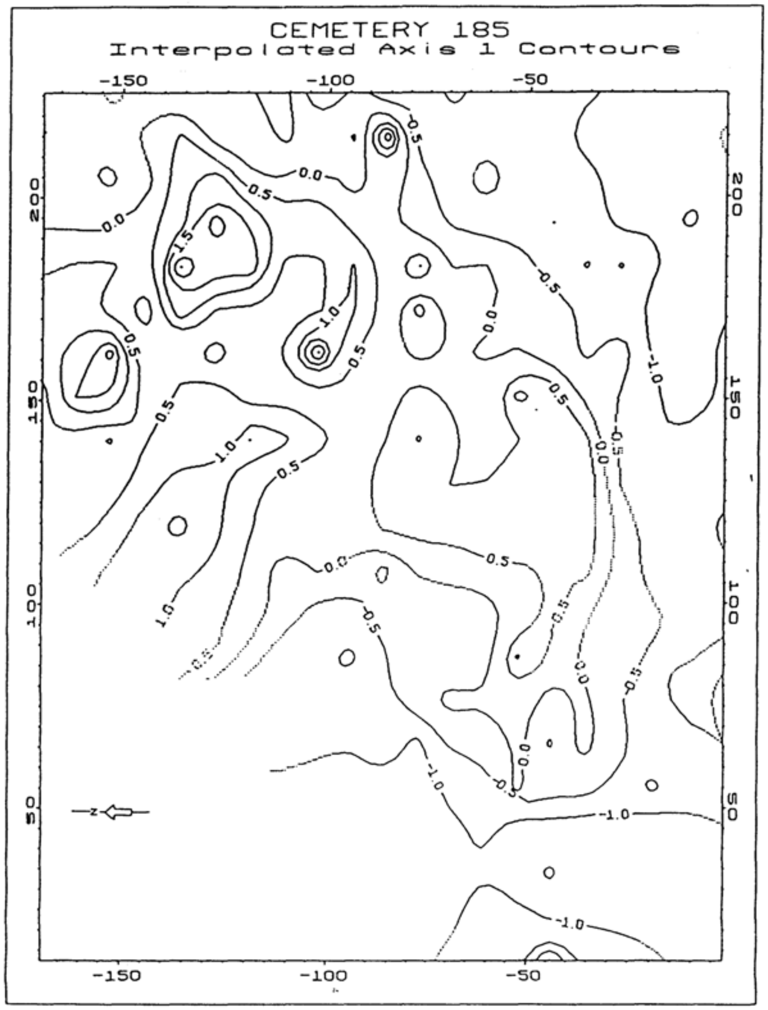


Figure 14.6 Contour map (Rockware—Gridzo) overlay of axis 1 (using the results from Wright) using the co-ordinates of the units.

the unit and also the finds. The range was 5 (high status) to 1 (low status). The criteria for this hierarchy of ranking are outlined below:

1. Simple tomb types (shaft, side niches and end niches) with no finds, other than a small number of beads or sherds of dubious origin. Excluded from the correspondence analysis due to lack of finds.
2.
    (a) Shaft tombs with the burial accompanied by kohl pot, perhaps wearing spiral earrings.
    (b) Simple tomb types (shaft, side niche, end niche) with one to four ceramic vessels.
    (c) Simple tomb types where the finds consist of a single seal and/or beads.
3. Simple tomb types with a large amount of pottery (more than four vessels) and/or other artefact types such as figurative pendants, kohl pots, metal vessels.
4. Chamber tombs in shafts and side niches and one end niche. Large amounts of pottery, and evidence of other 'high status' objects such as eg. a funerary mask.
5. Chamber tombs with 'ramps', mostly plundered. Large amounts of pottery and status objects such as funerary masks are deduced from the sophistication of the grave architecture.

The assignment of a status ranking to the individual graves involved a measure of subjectivity as well as a certain amount of presupposition as to the manner in which status or wealth is to be measured. In practical terms this means that the assumption was made that all of the examples of the relatively complex construction 'chamber with ramp' represented high status burials,

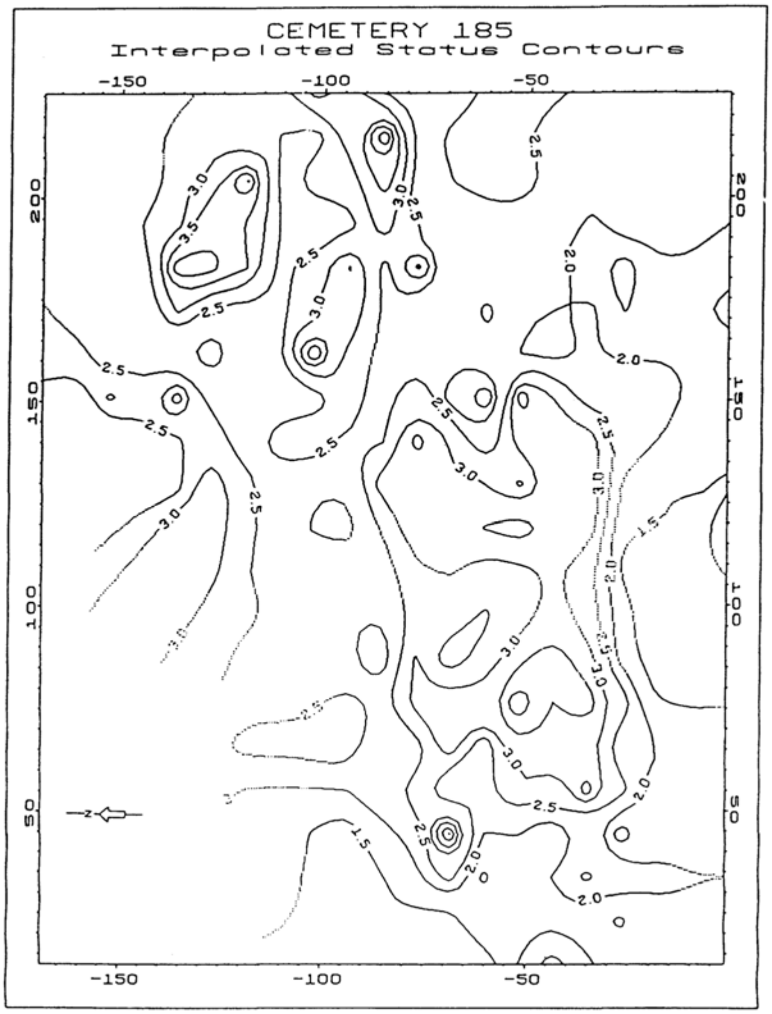


Figure 14.7 Contour map, overlay of status ranking using the co-ordinates of the units.

even though some were totally plundered and lacking in finds. The relative quantity of an object was used to suggest its level of availability and thus applied as a measure of its value as a status object. This meant that the relatively plentiful seals, when found as a single find, gave a lower status ranking ('2') than figurative pendants ('3') found in a much smaller number of units.

In this way each burial unit had a status ranking ascribed to it. In a way similar to that outlined above, an interpolated status contour map of the cemetery as a whole can be produced (Fig. 14.7). It should be noted that this depends upon treating the status variable as mathematically continuous and the results should accordingly be treated with some caution. It is however useful to compare the results of the contour map in Fig. 14.6 and the resulting overlaps which clearly indicate a correlation between status and temporal variability. For more detailed considerations of status variation in different time phases the more straightforward method of mapping the status distributions shown below is to be preferred.

Positing that the distribution of the burial units along the first axis does indeed represent a time trend, a series of figures (14.8–14.11), all drawn to the same scale, were produced to show the temporal variation in tomb type and status in different periods of cemetery use. The time units obtained from arbitrary cuts of the first axis of the correspondence analysis are likely to vary in length. An overall view of the results confirms the impression gained from the previous results, where the units and the attributes were projected onto two axes, of the progressive replacement of different tomb types.

All of the above analytical approaches encouraged a further reworking of the matrix and its results. A new version of the matrix was constructed, eliminating some of the more limited variables, such as construction details only relating to a small number of units and decoration details for a small number of specific ceramic subtypes. The choice of burial units was also reviewed and reduced to 474, excluding those where there could be some question as to whether the finds represented the original find repertoire of the unit. The new matrix was then run on the revised version of the CANOCO program. The eigen values for the second analysis were 0.505, 0.383, 0.337, 0.273, with a higher value for the first axis in the second analysis run with

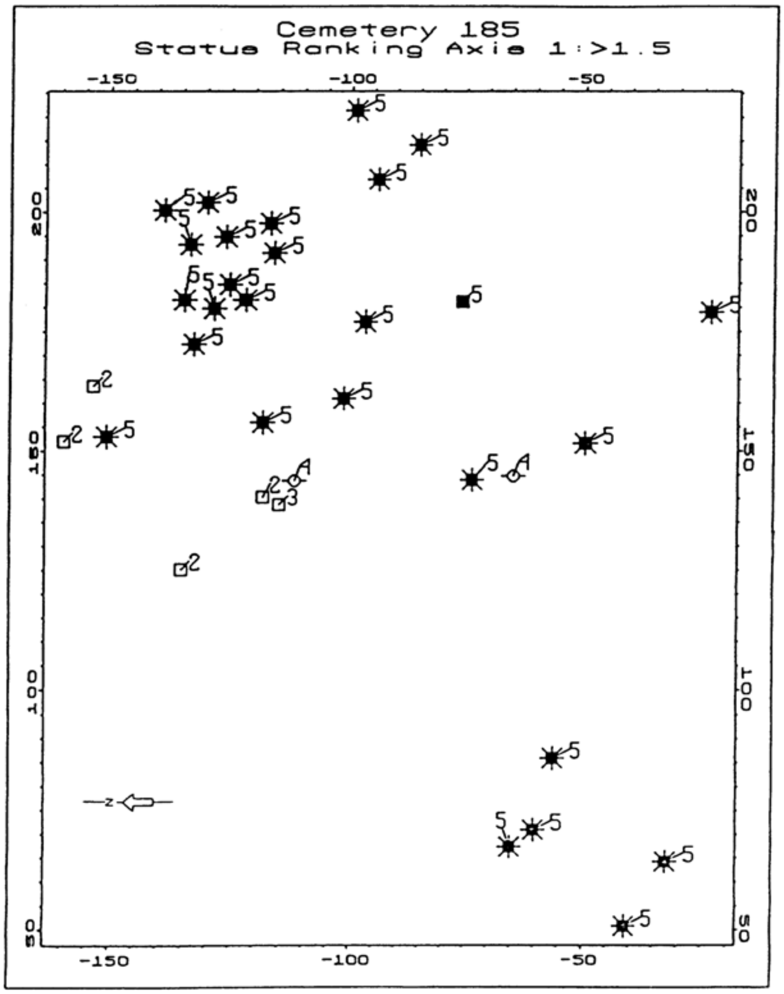


Figure 14.8 Units with a scoring on axis 1, higher than 1.5. The unit types are indicated using the symbols found on the key. The numbers refer to the status ranking of the individual units.

the CANOCO program, 0.505, as opposed to 0.410 for the preliminary run using the MV-Arch package. This suggested a more stringent seriation on the first axis using the edited version which became the basis for the final publication (Säve-Söderbergh & Troy 1991). It should be stressed however that the difference between the results of the two analyses was minimal as may be noted in the comparison of axis 1 and 2 of the two versions for some selected variables (Fig. 14.5).

Taking the results of the CANOCO run as the final version, the ranking of the units along the first axis provided by the CANOCO program was used as a seriation. The finds with relevance to dating were located along the first axis, using the score of the unit in which they were found. A division into three main phases, with seven subphases was made. The divisions were placed taking into consideration the occurrence of seals inscribed with royal names as *terminus post quem* indicators and to the shift in unit types, ceramic trends and find types. Parallels to the finds from this site, found elsewhere in Egypt and Nubia, were also regarded as significant in the placement of the lines of division.

Fadrus	Ia–Ib	Early 18th Dynasty, Pre-Hatshepsut c. 1570–1503 BC
	Ia	Axis 1—215120
	Ib	Axis 1—11995
Fadrus	IIa–IIc	Hatshepsut/Tuthmosis III-Tuthmosis IV c. 1503–1417 BC.
	IIa	Axis 1—94 51
	IIb	Axis 1—50 –36
	IIc	Axis 1— –37–102
Fadrus	IIIa–IIIb	Amenhotep III-Tutankhamun c. 1417–1352 BC.
	IIIa	Axis 1— –103–147
	IIIb	Axis 1— –147–200

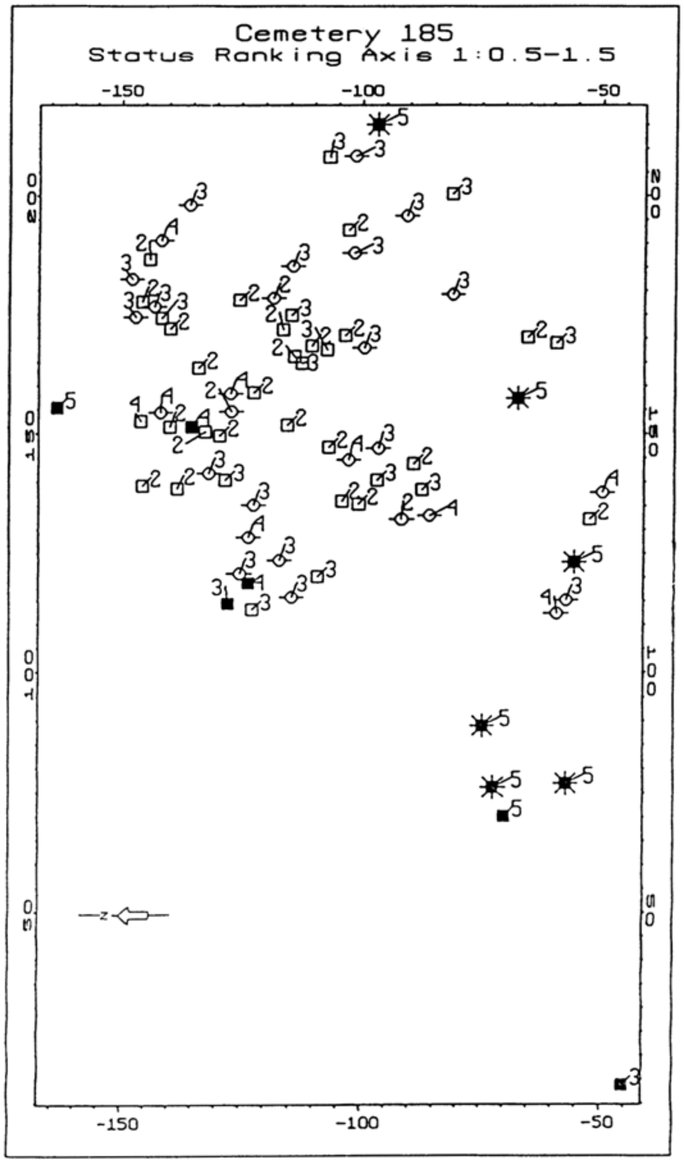


Figure 14.9 Units with a scoring on axis 1 in the interval 0.5–1.5.

These periods can, in relationship to Egypto-Nubian relations in Lower Nubia, be described as periods of the re-establishment of Egyptian hegemony (Fadrus I), the stable presence of Egyptian authority in a period of acculturation for the native Nubian communities (Fadrus II) and a period of decline, as the Egyptian presence diminishes and, it would appear, Lower Nubia is abandoned.

Treating the units and their contents as historical blocks facilitated a process of comparison whereby it was possible to see the variation of tomb types, ceramic combinations and find categories in relationship to the phases. The quantitative occurrence of the various ceramic types were also placed in a comparative framework using the seven subphases. Seen in relationship to the history of the area, however, it was perhaps of greater interest to observe the manner in which the relative proportion of high and low ranking units fluctuated during the different phases (Fig. 14.12). The percentage of each ranking group was calculated within each phase and this showed that while during the first phase the highest ranking group was around 20–25% of the units included in the analysis, this group diminishes in the second period and entirely disappears during the third phase. The lower ranking tombs, however, represented by about 36% in Fadrus Ia, constitute c. 91% of the units in Fadrus IIIb. In the middle period, during which time the Egyptian presence is consolidated in Lower Egypt and there is evidence of local Egyptianized traditions, it is possible to discern a more well differentiated hierarchy with a possible population increase in the form of the numerous lower level burials, homogeneous in both construction and funerary gifts. The general picture which the analysis of the phases presents coincides with the overall development of Lower Nubia, while offering new insights into the specific distribution of tomb types, artefact and pottery.

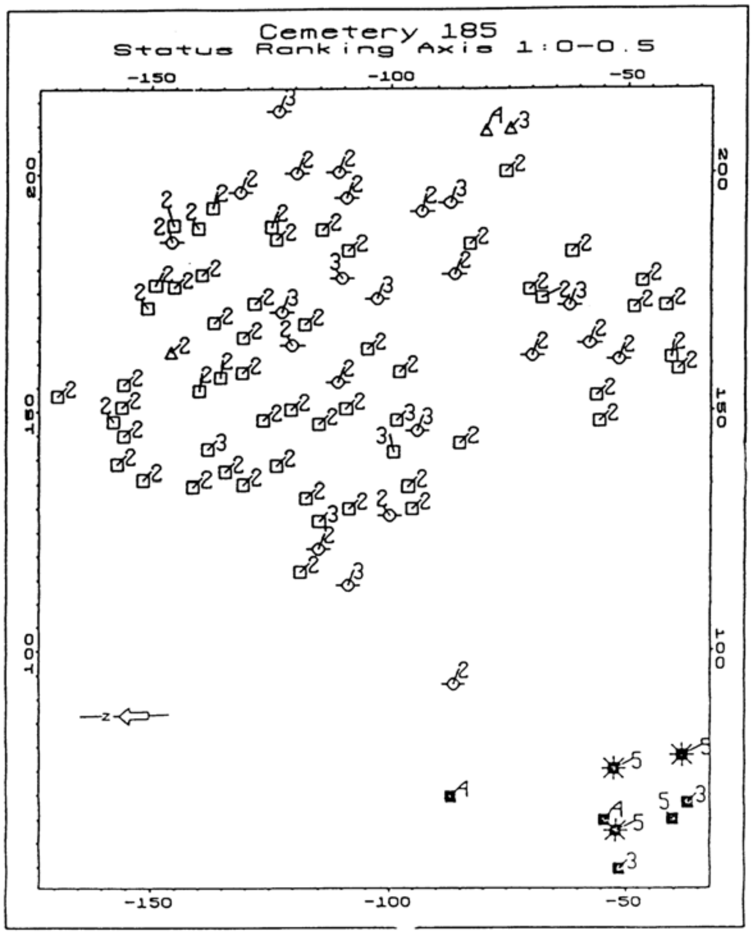


Figure 14.10 Units with a scoring on axis 1 in the interval 0.0–0.5.

The treatment of the units, using first a seriation along axis 1 and then a division into temporal blocks, gave a clear picture of the development of the site as a whole. An additional use of axis 1 was discovered in the examination of specific variables, whether tomb type, body position or find category. In each of the two analyses the variables were given scores using axes 1 and 2. These scores reflected the relative distribution of each variable among the units and, seriated, express the general status of each variable as ‘early’, ‘middle’ or ‘late’ within the context of the cemetery’s development. However, although this score can be regarded as an accurate ‘summing up’ of the chronological status of each variable it does not provide any indication of the chronological range of distribution. When however the artefact is mapped along the first axis using the score given the unit in which it is found the illustration of such a distribution can be obtained (Fig. 14.13). This was done with the most important ceramic types, providing, once again, an easily interpreted pictorial representation of the gradual introduction and eventual decline of the many different ceramic types found on the Fadrus site.

Many different analytical aids have been applied to the documentation of site 185: database analysis, correspondence analysis, digitization of co-ordinates and contour mapping using a combination of the digitization and the results of the correspondence analysis. The results which have been obtained constitute an exciting addition to our knowledge of New Kingdom Lower Nubia.

Closing remarks

The international campaign to save Nubia, of which the Scandinavian Joint Expedition to Sudanese Nubia was a part, produced significant quantities of archaeological data. In that respect it vies with the period 1880–1920 during which the most intensive excavation of the Nile Valley took place. As the onerous responsibility of basic description of irreplaceable finds imposed by the rescue nature of the campaigns is met, it is important not to lose sight of the analytical potential of the rich and complex nature of this material.

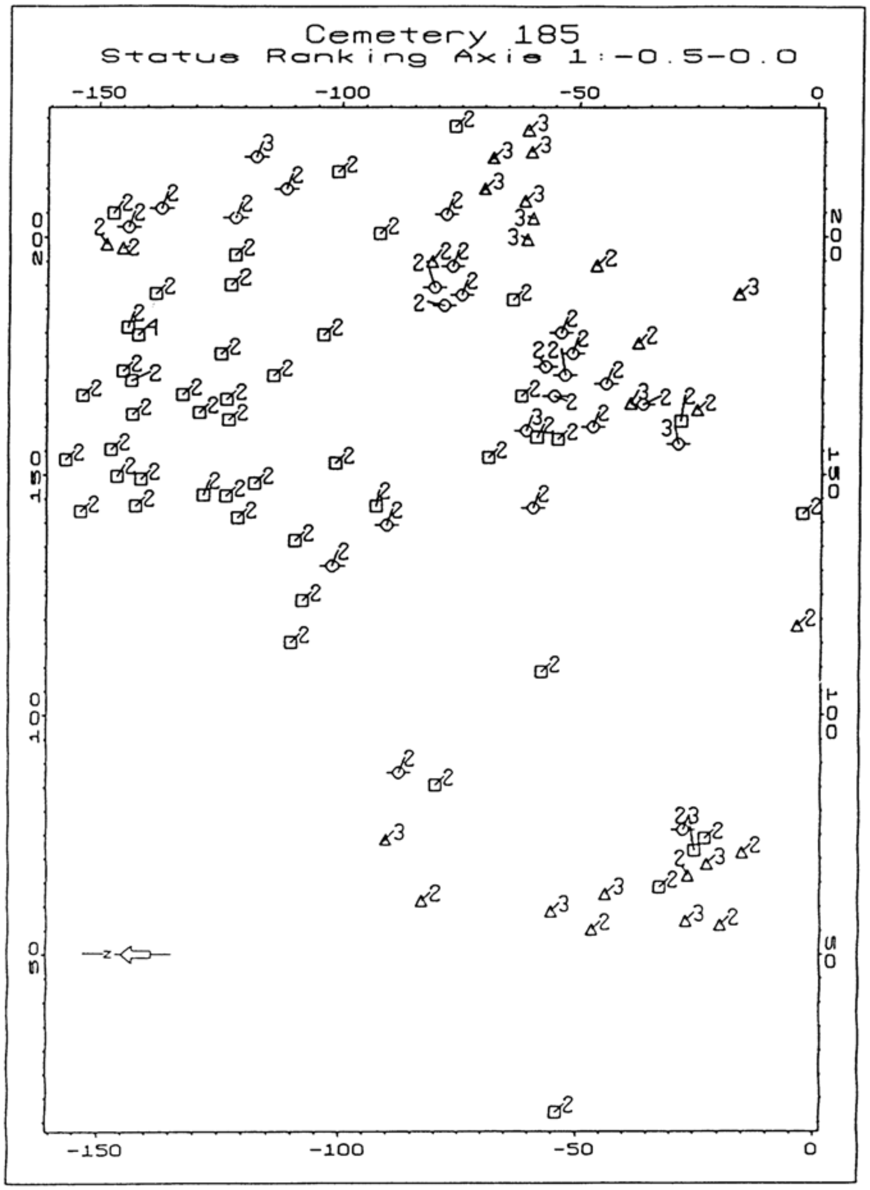


Figure 14.11 Units with a scoring on axis 1 in the interval −0.5–0.0.

Acknowledgements

Sincere thanks are due to Prof. T.Säve-Söderbergh, Prof. R.Holthoer and Dr T.Madsen for stimulating comment in the course of production of this chapter.

References

Adams, W.Y. 1977. *Nubia: corridor to Africa.* London: Allen Lane.

Benzecri, J.P. 1973. *L'Analyse des Données, I Le Taxonomie, II L'Analyse de Correspondances.* Paris: Paris du Nord.

Bietak, M. 1966. *Ausgrabungen in Sayala-Nubien 1961–1965. Denkmäler der C-Gruppe und der Pan-gräber-Kultur.* Wien: Österreichische Akademie der Wissenschaften. Denkschriften Band 92.

Bietak, M. 1968. *Studien zur Chronologie der Nubischen C-Gruppe. Ein Beirtrag zur Frühgeschichte Unternubiens zwischen 2200 und 1550 vor Chr.* Wien: Österreichische Akademie der Wissenschaften. Denkschriften Band. 97.

Bietak, M. 1987. The C-Group and Pan-Grave Culture. In *Nubian Culture Past and Present,* Hägg, T. (ed.), 113–128. Stockholm: Almqvist & Wiksell International.

Bolviken, E., E.Helskog, K.Helskog, I.M.Holm-Olsen, L.Solheim, & R.Bertelsen 1982. Correspondence Analysis: an alternative to principal components. *World Archeology* **14(1),** 41–60.

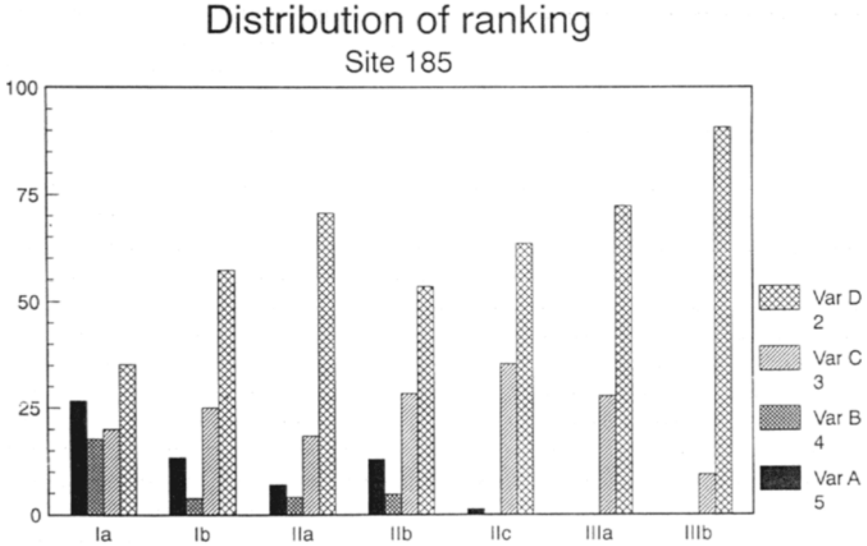


Figure 14.12 Diagram illustrating the differentiation in status in the subphases.

Bonnet, C. 1978. Fouilles archéoloques à Kerma (Soudan): rapport préliminaire de la compagne 1977–1978. *Genava* **26,** 107–127.

Bonnet, C. 1980. Fouilles archéoloques à Kerma (Soudan). *Genava* **28,** 31–72.

Bonnet, C. 1982. Fouilles archéoloques à Kerma (Soudan). *Genava* **30,** 29–53.

Bonnet, C. 1984. Fouilles archéoloques à Kerma (Soudan). *Genava* **32,** 5–20.

Bonnet, C. 1986. *Kerma, Territoire et Metropole.* Cairo: Institut Française d'Archeologie Orientale du Caire.

Gratien, B. 1978. *Les cultures Kerma. Essai de classification.* L'Université de Lille III: Villeneuve— d'Ascov.

Kemp, B. 1976. Dating pharaonic cemeteries. Part I: non-mechanical approaches to seriation. *Mitteilungen des Deutschen Archaeologischen Instituts Abteilung Kairo* **31(2),** 93–116.

Säve-Söderbergh, T. 1941. *Ägypten und Nubien: ein beitrag zur Geschichte Itagyptishe Aussen politik.* Lund: Hakan Ohlssons Boktryckeri.

Säve-Söderbergh, T. 1963. Preliminary report of the Scandinavian Joint Expedition: archaeological investigations between Faras and Gemai. *Kush* **XI,** 47–69.

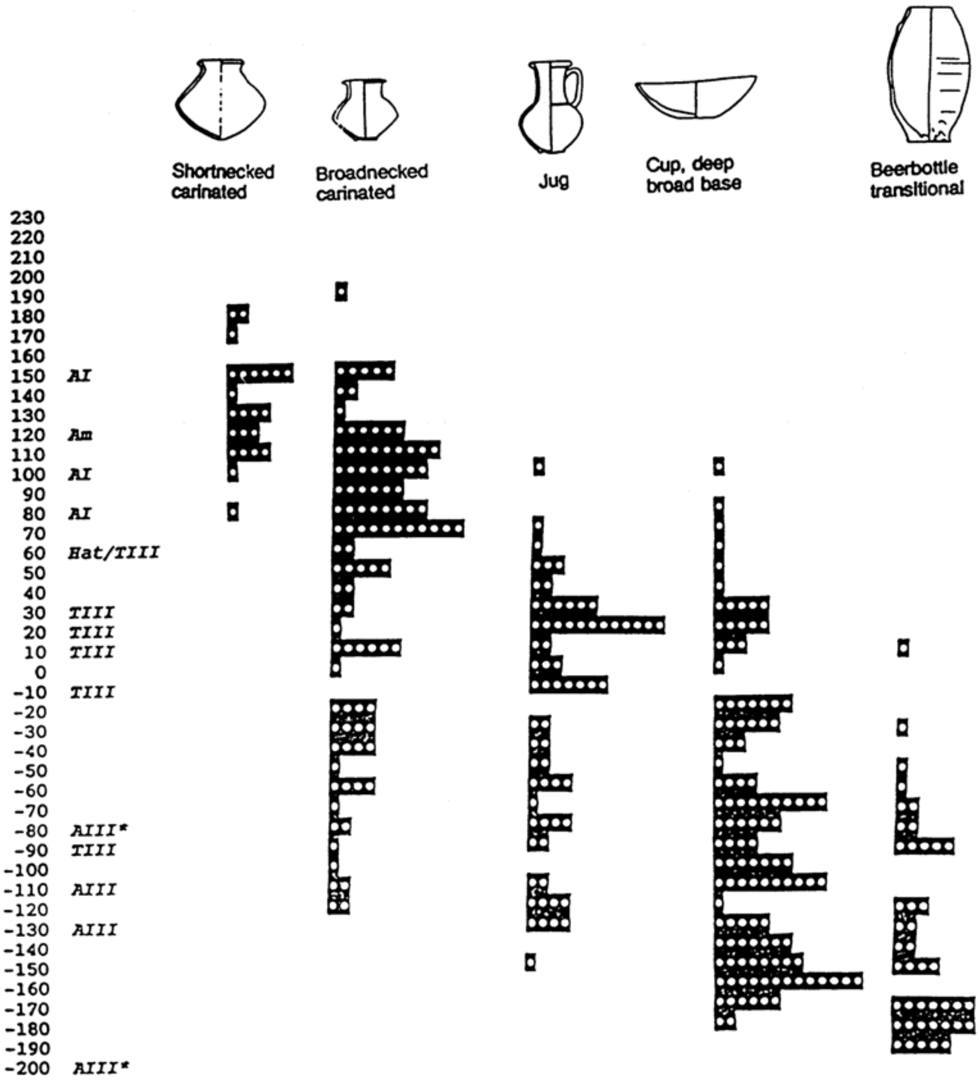
Säve-Söderbergh, T. (ed.) 1987. *Temples and Tombs of Ancient Nubia. The International Rescue Campaign at Abu Simbel, Philae and other Sites.* London: Thames & Hudson.

Säve-Söderbergh, T. & L.Troy 1991. New Kingdom Pharaonic Sites. *Swedish Joint Expedition* **5** (2–3).

Sinclair, P.J.J. & L.Troy 1988. Organizing archaeological data from the Scandinavian Joint Expedition: Preliminary results of a multivariate analysis. *Informatique et Egyptologie* **5,** 43–45.

Sinclair, P.J.J. & L.Troy 1989. A multivariate analysis. In *Middle Nubian Sites,* Sävë-Söderbergh, T. (ed.), 273–290. Partille, Sweden: Paul Osröm Editions.

TerBraak, C.J.F. 1988. *CANOCO. A FORTRAN program for canonical community ordinations by [partial] [detrended] correspondance analysis, principle components analysis and redundancy analysis (Version 2.1).* Technical Report LWA-88–02. Wageningen, Netherlands: Agricultural Mathematics Group.



* in secondary context

Am > Seals inscribed Nebpehtyre (Amosis, ca 1567 – 1546 BC)

AI > Seals incribed Amenhotep or Djesarkare (Amenhotep I, ca 1546 – 1526 BC)

Hat/TIII > Unit containing seals inscribed Hatshepsut and Menkheperre (Hatshepsut/Tuthmosis III ca 1503 – 1482 BC)

TIII > Seals inscribed Menkheperre (Tuthmosis III, ca 1504 – 1450 BC)

AIII > Seals inscribed Nebmaatre (Amenhotep III, ca 1417 – 1379 BC), note two examples not in situ

Figure 14.13 Distribution of selected ceramic types using axis 1.

15

Cultural change, the prehistoric mind and archaeological simulations

MARTIN BISKOWSKI

Introduction

Individual decision making and the transmission of cultural meanings, symbolism, and ideology have all influenced the evolution of prehistoric societies, but these processes generally leave behind relatively poor archaeological evidence. The systematic manner in which members of a culture responded to and transmitted ideas and information must have placed important limits on the direction and speed of cultural change. Consequently, the means of information processing employed within cultures presumably influenced significantly the course of cultural evolution. Unfortunately, archaeologists often can do little more than record highly equivocal evidence about the presence and nature of human information processing mechanisms in past societies. Obviously, archaeologists and cultural anthropologists generally cannot study directly the long term operation of various information-related mechanisms. Even for short term ethnographic research, the operation of these mechanisms is fairly complex and difficult to observe. Naturally, tracing the consequences of these mechanisms over lengthy periods adds further complications.

One avenue for studying the implications of long-term information-processing is through computer simulation. Researchers in recent years (eg. Boyd & Richerson 1985; Cavalli-Sforza & Feldman 1981) have shown how the transmission of cultural information via such widely varied channels as verbal communication, role models, and passive symbol systems may be described and operationalized in formal mathematical models. Before information processing simulations can be integrated into either practical studies of the past or possible explanations of what we can observe about the past, several practical problems must be resolved. In particular, one key problem is to create a generalized simulation structure in which the processes by which cultural information is communicated and utilized by various actors are clear to both the simulator and the simulation's intended audience.

One promising way to increase the comprehensibility of such simulations involves the judicious use of artificial intelligence (AI) representations of knowledge. This requires that conclusions be reached from information supplied to the actors developed by the simulation. Also, the simulator must provide for the management of events within the simulated world based on different possible contingencies. In both cases, the production-rule structure of some expert systems can provide a clear, explicit outline of how different possible conclusions are reached, sidestepping cryptic, often unpersuasive simulation procedures. Moreover, an expert system structure provides a representation which accurately portrays many of the subtleties and the effectiveness of cultural reasoning (cf. D'Andrade 1981; Read & Behrens 1989; D'Andrade 1989).

General goals

Recent efforts by postprocessualists (eg. Hodder 1982a; Hodder 1982b; Miller & Tilley 1984; Hodder 1985; Hodder 1986) have alerted archaeologists to the relationship between how past individuals viewed their world and the archaeological record. However, archaeologists' efforts to pursue the postprocessual program have encountered numerous problems (Earle & Preucel 1987; Chippindale 1987; Renfrew 1987). Until recently archaeologists tended to ignore cognitive phenomena because making inferences about past mental processes is always difficult even when it may be possible. They have tended to simplify the role of past human cognitive processes by using two basic assumptions about how individuals make decisions and adopt new ideas.

The first assumption is that individuals within a society will act according to their perceived best interests. This economic rationalism may be constrained by the cultural and social perspective of the individual (cf. Limp & Carr 1985). The second, often unstated, assumption is that behaviours or strategies which are advantageous will, over archaeological time spans, become so apparent that everyone will quickly adopt them (compare with the concept of direct bias transmission used by Boyd & Richerson 1985, pp. 10, 35) or else the advantage of the new behaviours or strategies will be so profound that a form of cultural group selection will take place, favouring those who have assimilated the new ideas.

While these assumptions cannot account for many kinds of contemporary cultural change, nonetheless they have been indispensable to archaeological analyses because they have allowed researchers to establish reasonable, loosely fitting boundaries on the probable ranges of past human behaviours. Archaeologists can detect gross changes in climate, resource availability, population size, subsistence techniques, settlement organization, apparent socioeconomic differentiation, and exchange patterns. Defining relationships between these data and the changing marginal costs of different resource procurement strategies is a relatively straightforward enterprise (eg. Earle 1980; Christenson 1980) which in many cases captures the essential rules structuring past situations of cultural change.

Some researchers, recognizing that archaeological research has progressed effectively without elaborate consideration of the way people actually acquire cultural knowledge or how they use it and other information to make decisions, may argue that archaeologists need not be concerned with ephemeral, cognitive phenomena. Unfortunately, the mechanisms which induce and influence cultural change do not exist entirely outside of individual human minds. Consequently, by ignoring cognitive phenomena, archaeologists risk considerable self-deception. Read (1985; in press) has illustrated some of the problems which may arise when we use formal models to describe archaeological phenomena which are not concordant with the processes generating that phenomena. We can generate plausible explanations of cultural change without considering the role which ‘invisible’ factors related to cognition might play. But, by ignoring such factors, we lose sight of important alternative explanations and therefore fail to evaluate properly non-cognitive explanations within the context of all reasonable alternative explanations.

One example of how cognitive phenomena may influence independently the rate and direction of cultural change comes from Reynolds’s (1986; see also Flannery 1986b; Flannery 1986a) simulation of the adoption of agriculture in Oaxaca. Reynolds’s simulation attempted to address the role of factors such as:

1. the ability of each member to collect and process information about the resource distribution.
2. the extent to which information is shared among the members.
3. the specific sets of decisions available to each member.
4. the way in which the individual decisions are integrated to produce a group decision.

The simulation that Reynolds devised contains only very simple cognitive phenomena—the passage of information and the determination of subsistence strategies. Nonetheless, the simulation trials produced a number of surprising results which can be traced directly to the cognitive elements of the simulation model.

Improvements in foraging efficiency, for example, took place most quickly under conditions of environmental variability which incorporated not only stressful periods (bad years) but also sufficiently nonstressful periods (good years) to allow for experimentation with, and evaluation of, new subsistence strategies. When provided with a uniform sequence of ‘average’ years, the simulated foraging group’s strategies did not converge on greater efficiency. Instead of optimizing subsistence strategies in the face of an unchanging environment, the group moved ‘aimlessly’ from strategy to strategy. This led Flannery (1986a) to underscore the importance of information availability to the process of adaptation. Also, the simulation results indicate that the process of coping with inconsistent information from highly variable environments can speed change towards agricultural practices independent of overpopulation, intergroup competition, and other, noncognitive factors commonly invoked in archaeological explanations.

The complexities of learning and communication provide only one set of cognitive influences on the dynamics of cultural systems. Another set of influences is the inertia of systems of cultural knowledge. One illustrative example is Conrad & Demarest’s (1984) explanation of the collapse of the Aztec and Inka empires. Although the growth of these empires was fueled by an expansionist ideology, later leaders apparently realized that continued expansion without adequate consolidation would not help but harm the respective states. Conrad & Demarest do not ignore the existence of entrenched social groups whose most immediate best interests depended on continued expansion (Conrad & Demarest 1984, p. 182), but the heart of their explanation contends that attempts to halt imperial expansion were blocked because the idea that expansion might be harmful could not be incorporated into cultural schemata which so actively emphasized the positive value of conquest and expansion.

Biological constraints on human cognition can also influence the evolution of cultures. The most illustrative example in the archaeological literature comes from Gregory Johnson’s studies of the relationship between human abilities to process information, the organization of administrative hierarchies, and the ability of such hierarchies in early states to maintain control over territory and enterprises (Johnson 1973, p. 153; Wright & Johnson 1975; Johnson 1982; Johnson 1988). Johnson argues that humans can only monitor and process a certain amount of information at a time. More precisely, humans, while thinking, can only keep track of around 6–7 pieces of information at the same time (see Miller 1956). Therefore, a bureaucrat in an early state administration directly responsible for monitoring and planning for 24 villages would encounter difficulties directing those villages towards some central purpose that someone administering to four districts of 6 villages would not experience. Stable administrative structure requires sufficient hierarchization so that the burdens on individual decision makers do not greatly exceed human cognitive limits. Otherwise, an administrative structure would lose control over whatever it was supposed to manage. This sort of cognitive limitation on the number of categories of items a person can think about at once should affect not just bureaucracies, but also should influence processes of thought in most domains of knowledge.

These three classes of cognitive phenomena—learning and communication of ideas, the inertia of systems of cultural ideas, and the biological constraints on human thought—all may impact the direction and speed of cultural change. Moreover, the existence of each of these kinds of influences on cultural change has long been recognized. But, in recent years, our understanding of the past has progressed to the point where issues involving cognitive phenomena are becoming unavoidable. While simulations modeling these cognitive phenomena cannot possibly duplicate all of the associated complexities, they still offer the opportunity of formalizing and operationalizing what up until now have been concepts restricted to verbal expression in explanatory models. By constructing these simulations, researchers may better evaluate models purporting to explain archaeological evidence and consider more directly the conditions under which cognitive influences on cultural change provide more plausible alternative explanations.

General problems in representation

Clearly, exploring these possibilities through simulation requires the development of some systematic means of modeling different cognitive phenomena. In particular, for these simulations to provide useful models of long term change they must model how these phenomena were translated into conscious decision making and other bases for human action which then, in turn, would influence the formation of the archaeological record. Reynolds's simulation provides one example of how learning and communication within human groups can be explored. Further inspiration might be sought in models of the transmission of cultural ideas developed by Cavalli-Sforza & Feldman (1981) and Boyd & Richerson (1985). Learning and communication, considered merely as social phenomena dependent on personal interactions and social relationships, are relatively simple to model. The transmission of cultural information can be treated similarly to the transmission of genes, usually with little consideration of the resistance to the reception of new ideas once they are transmitted.

In contrast, modeling the inertia of cultural knowledge discussed earlier in the description of Conrad and Demarest's work causes considerably more problems because the model requires a more sophisticated treatment of the reception of new ideas. The 'inertia' represents the difficulty or facility with which members of a cultural group might accept specific, new ideas because of preexisting cultural ideas. Accordingly, resistance to change might develop when a new idea conflicts with one or more of the basic concepts of an established system of cultural ideas. As an example, a new method of foraging might be adopted more readily by hunting-gathering bands if it fits easily into established notions of foraging activity and social interactions than otherwise. A model realistically reflecting inertia must contain some procedural representation of an individual's knowledge and some measure of how conflicts with preexisting cultural (or personal) knowledge might limit an individual's receptivity to new ideas. In the representation of preexisting knowledge, both the content and the structure are important. The content of preexisting knowledge may contradict a new idea, and the importance of a particular knowledge rule within the structural representation should indicate how much disruption might be caused by replacing it with a new rule of different content.

Simulating the biological limitations imposed on human cognitive abilities which concerned Johnson similarly calls for some structural representation of human thought processes. Here, however, we are concerned with limiting the size and shape of structural representations of knowledge irrespective of the content. For the administrator dealing with 24 villages, unless these villages are otherwise distinguished, his or her decision procedures concerning any specific village will be applicable to any other village posing similar problems. Within the more hierarchically organized bureaucracy, decision procedures may be directed more specifically either at individual villages or at groups of villages.

Expert systems in simulations

One way of incorporating these features into a simulation is to first define different kinds of individual behavioural perspectives and then use artificial intelligence concepts to model how different individuals will behave within a simulation. Different researchers (eg. Reynolds & Zeigler 1979; Doran 1982; Reynolds 1986; Doran 1987; Renfrew 1987) have promoted the use of artificial intelligence concepts for modeling prehistory. Doran proposed the use of multi-actor models wherein the actions of each actor are determined by a goal (energy acquisition), a 'knowledge base', and alternative plans for achieving that goal derived from the knowledge base. Renfrew expanded on this conceptualization considerably by adding the acquisition of prestige, participation in supernatural rituals, and other goals embedded in particular cultural contexts.

These conceptualizations of how simulations of human cognition and communication should be constructed to assist archaeological research indicate that expert systems models can play an important role in such simulations. An expert system contains a procedural representation of a body of knowledge which can be used to anticipate an individual's reasoning

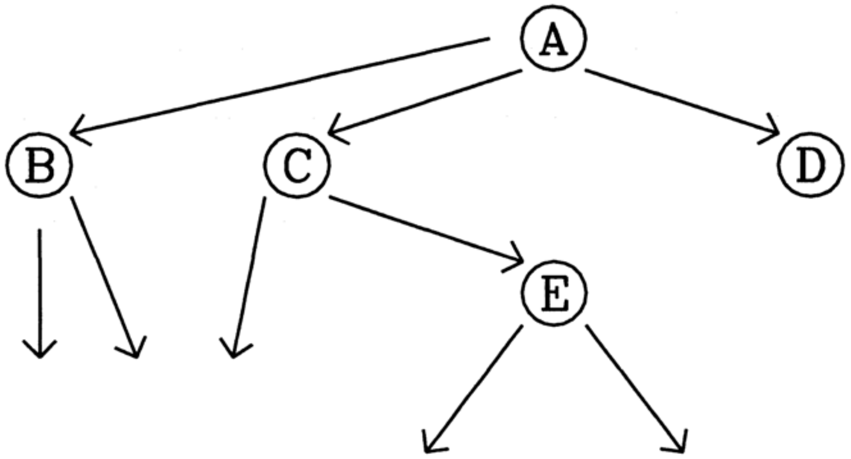


Figure 15.1 Reference paths among rules A-E.

processes. Generally, the knowledge representation is encoded in a corpus of individual production rules which govern how various predefined conclusions are to be reached. Characteristically, expert systems reach conclusions through the recognition of extensive sets of clues in reaching a conclusion rather than through exhaustive searching (Simon & Kaplan 1989). Ideally, a simulation of human cognition should mimic the intuitive nature of cultural reasoning and the misfirings of unschematized reasoning (cf. D'Andrade 1989), and an expert system format will allow a realistic imitation of these influences on an individual's cognitive performance.

Expert systems have been developed to simulate human performance in medical diagnosis (Clancey 1983; Politakis & Weiss 1984), cotton crop management (Lemmon 1986), and other highly specialized areas of knowledge. Expert systems are extremely useful for modeling the kinds of reasoning processes expressed in folk knowledge (Read & Behrens 1989), and, importantly, this usage can be extended to include cultural or personal reasoning processes about the 'proper' methods of securing some goal, such as how to obtain important goods through exchange.

Normally, the development of an expert system is influenced by those factors which more generally affect the elicitation of folk knowledge (Cicourel 1986). The methodological problems of eliciting accurate information from living informants do not apply to the construction of models of the distant past because all reasonable informants are long dead. Instead, by considering the cultural, social, and economic contexts, and also by applying ethnographic and historic analogies, a simulator can construct one or more plausible simulations of past cognition. Naturally, this approach may require the simulator to consider a broad range of plausible models to determine whether there are substantive differences between them.

Expert systems provide a means for representing the inertia of a system of cultural ideas. While expert systems reach conclusions using a number of different pieces of information, some are more important than others within a given reasoning process. Consider the prehistoric 'expert' partially modeled below and in Figure 15.1.

RULES:
A. Can obtain utilitarian goods through exchange IF
1. Possesses exchangeable excess goods
AND 2. Appropriate exchange partners are available
AND 3. Can negotiate exchange
B. Possesses exchangeable excess goods IF
1. Some household products/possessions exceed forseeable needs
AND 2. Excess goods are exchangeable
C. Appropriate exchange partners are available IF

1. Knows candidates for exchange partners
AND 2. Some candidates are socially appropriate
D. Can negotiate exchange
1.…
AND 2.…
E. Some candidates are socially appropriate IF

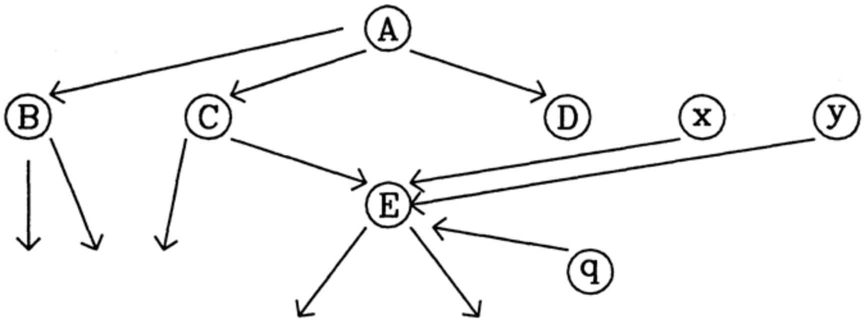


Figure 15.2 Reference paths among rules A-E plus other rules referencing rule E.

1. Some candidates are members of extended kin group OR
AND 2. Some candidates are neighbors of similar status
F…., etc.

Figure 15.1 expresses the structural relationship between different rules, with each arrow indicating that one rule references either another rule or a directly evaluable fact. Since this example only includes a small, simple subset of the knowledge structure, there is little that is remarkable about the structure or position of any particular rule.

Nonetheless, the simple structural relationships between rules as expressed in Figure 15.1 provide one way of modeling the resistance, or inertia, of any given rule to change. Rule B determines whether the actor possesses excess goods which can be exchanged for other goods. Substituting a new rule for B implies the elimination and addition of rules referenced below B. At the same time, putting a new rule in the place of B may change the performance of A and of other rules which might reference A. The amount of inertia resisting a rule substitution might, accordingly, be modeled by the number of rules whose content are changed by the substitution and the number of rules which are no longer referenced.

The effectiveness of this way of modeling inertia can be seen more easily by broadening the consideration of the above example to include other aspects of an expert's knowledge. In many of the expert's other social interactions the expert also must find appropriate persons with whom to interact. The additional arrows drawn to rule E in Figure 15.2 reflect that rule E is referenced by a number of rules outside of the immediate context of exchanging utilitarian goods. Consequently, changing the content of rule E necessarily changes the content of several rules. In this sense, rule E possesses greater inertia than rule B. In much the same way, a rule governing social interaction which permeates many different domains of cultural knowledge should be more difficult to displace than a rather specific rule about whether a household possesses excess exchangeable goods.

Human cognitive limitations can also be modeled within an expert systems format by placing certain kinds of limits on the structure and size of the knowledge representation. One might impose an upper bound on the number of conditional statements defined in each rule. Also, in keeping with Johnson's ideas, one might also limit how many different categories of an item are distinguished by conditional statements. So, the conditional clauses used by an expert administrator dealing with six villages might include tailored considerations of one or more villages:

Rule
J…. IF
1. Village 5's crops are below normal

But an expert administrator dealing with 24 villages might be limited to more general conditional clauses, such as:

Rule
J…. IF
1. A village's crops are below normal

As a more general observation, the structural relationships between rules within an expert system promote the use of a variety of different descriptive concepts from mathematical graph theory, organization theory, and systems theory. The concept of centrality could be applied to Rule E in Figure 15.2 because of its importance in the evaluation of several rules. Concepts like reducibility and redundancy can be applied to knowledge structures to describe structures where certain rules are emphasized by redundant references to them. For example, in order to reach a particular conclusion, a given expert system in an irreducible form will evaluate a given rule no more than once (Fig. 15.3).

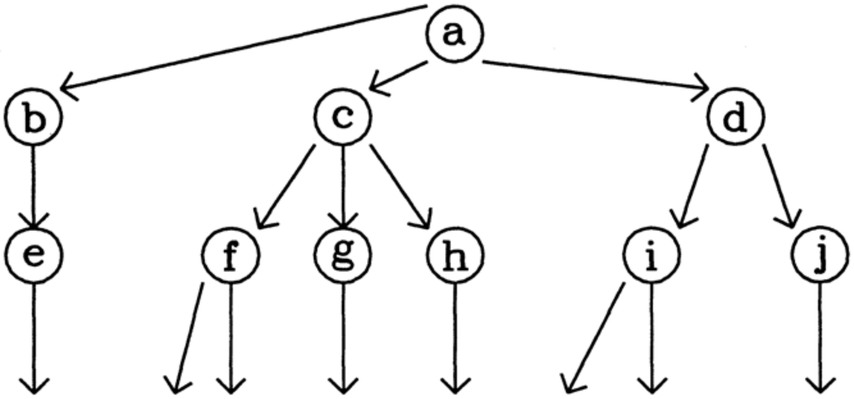


Figure 15.3 Irreducible reference path map (rules b-j are referenced only once).

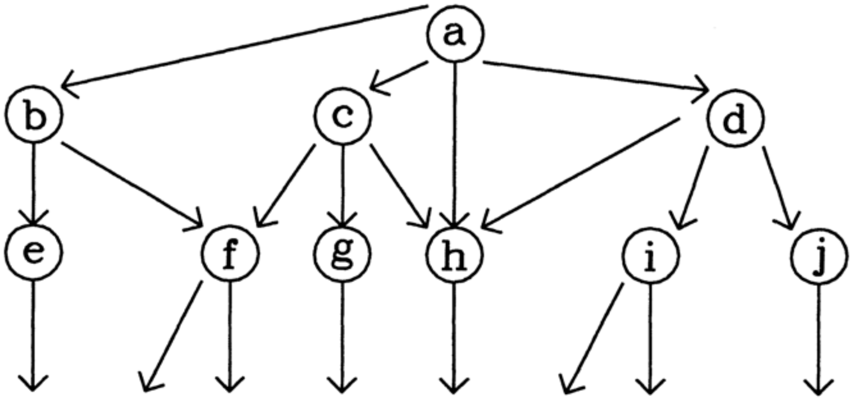


Figure 15.4 Reducible reference path map (several redundant references).

But when several different rules in the path towards a particular conclusion each refer to yet another rule as a precondition (Fig. 15.4), the structure is reducible in the sense that one or more of these references might be eliminated without changing the conclusions reached by this part of the expert system in response to any set of information. The knowledge structure specifies that a specific condition must be met at several different points in the artificial reasoning process. Yet, after the first time that specific rule is evaluated, further, redundant evaluations of the same rule will change nothing.

The concept of redundancy is particularly useful for understanding and modeling inertia within the context of expert systems. Another way to model inertia is to assign values to each rule within an expert system according to the frequency of times other rules reference it. Larger values can represent rules more difficult to remove or replace. In this model each redundant reference to a rule will increase the inertia of that rule. In a culture where most interactions depend on kin memberships, for example, many evaluation procedures will contain redundant references to rules defining kin membership. Naturally, given their central importance, such rules should possess considerable inertia.

In sum, expert systems are only synchronic representations which model how simulated individuals should respond when given certain goals and a set of information. Within a simulation, the operation of an expert system provides a useful basis for determining whether a given individual is successful in attaining his or her goals at a specific point in time.

Unlike human experts, however, expert systems do not revise themselves. Within a simulation, such revision must be generated elsewhere. In situations where ideas and concepts can be modeled satisfactorily as traits which can be transmitted independently of each other, the use of expert systems provides no special advantages. But in situations where the interrelationships between different concepts are important in modeling the dynamics of their transmission, the structure of an expert system provides a basis for modeling these interrelationships and the extent to which these interrelationships influence the reception of new or alternative ideas.

Clearly, the methods by which the structures of expert systems can be used for modeling the transmission of cultural ideas need further exploration and refinement. The main point of this discussion is that expert systems potentially can provide archaeologists and other researchers interested in cultural change with the means to describe phenomena like inertia within simulation models. By transforming explanations of cultural change posed largely as verbal conjecture into formal simulations, we can explore the implications of such explanations much more readily and, hopefully, evaluate their worth within the context of alternative explanations.

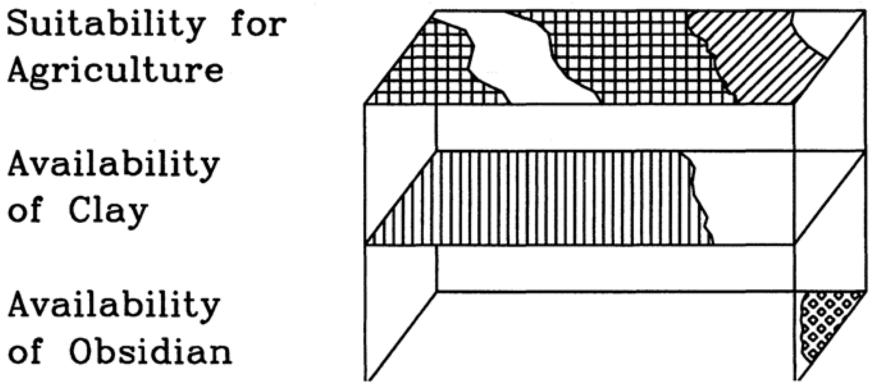


Figure 15.5 Resource surface.

Outline of a design

Expert systems will provide two general functions within the simulation design. First, an expert system can be used to model how an individual will respond to the range of possible situations provided in the simulation. Thus, a single expert system can be assigned to model how a large class of individuals will perform, even though each individual will have a different perspective and each as a result will respond differently within the simulation. Moreover, the rules governing different responses can be modified easily, so new kinds of individuals can be adapted from the old.

Second, expert systems will manage the simulation. The procedures used for revising the representations of different individuals' knowledge are especially important and must be clear and understandable to both the simulator and anyone he/she seeks to persuade with the simulation. Also, the simulated individuals must be provided with the goals and stimuli which spur their information processing. In essence, a world of resources, social networks, and various goals must be provided for the simulated actors. For my research purposes certain consequences of individual interactions— the acquisition of exchange partners and the recognizable and objective success of exchanges—must be resolved. The use of expert systems for managerial purposes is not simply gratuitous but is, instead, in line with the more normal functions of expert systems. The manner in which the simulation performs its task must be left open to scrutiny and criticism. The managerial systems are intended to simulate the simulator's own expertise and reasoning (within limits) as an explicit set of production rules subject to individual examination (cf. Clancey 1983).

As a specific example, the simulation which I am developing contains three major components. One is a resource surface which models the areal distribution of cultivable land and other natural or artificial resources in space. The second is a set of decision-making subprograms which generates the reasoning of different simulated individuals. The third component is a set of referee subprograms that manage where decision-makers are, what kinds of information are available to each, and how decisions and nondecisions are translated into modifications of, and control over subsets of the resource surface.

The resource surface

The resource surface represents the regional distribution of different resources important to early complex societies (Fig. 15.5). As a way to simplify my own process of conceptualization in designing the surface, I principally have made use of resources common to highland Mesoamerica. Resources incorporated into the surface include land with different potentials for dry farming and more intensive forms of agriculture, obsidian, clay, timber, and basalt. The resource surface presents the spatial relationship amongst these resources in a manner which will have to be resolved by the simulated individuals.

Different researchers of Mesoamerican prehistory (eg. Sanders & Santley 1983; Hassig 1985) have investigated the costliness of meeting supply requirements through exchange at various distances, and these direct economic consequences of the resource surface will be reflected in the other components of the simulation. Also, the feasibility of controlling different kinds of resources will be a factor considered by the AI portions of the simulation.

Decision making modules

The decision making modules are modeled as expert systems which simulate an individual's cultural and personal knowledge of different methods of interacting with the resource surface in order to fulfill various goals and needs (Fig. 15.6). Rules for selecting exchange partners and the values of various items are also included in these modules. Presently, the social position of this individual has not been defined beyond being someone with the power to act in exchange relationships on behalf of some corporate group. Within the program architecture, the production rules used by a given simulated individual will be drawn

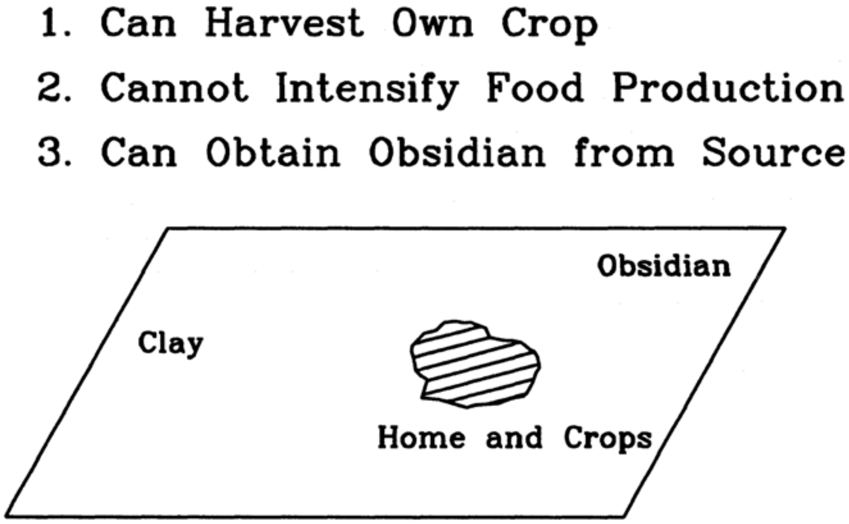


Figure 15.6 Actor j.

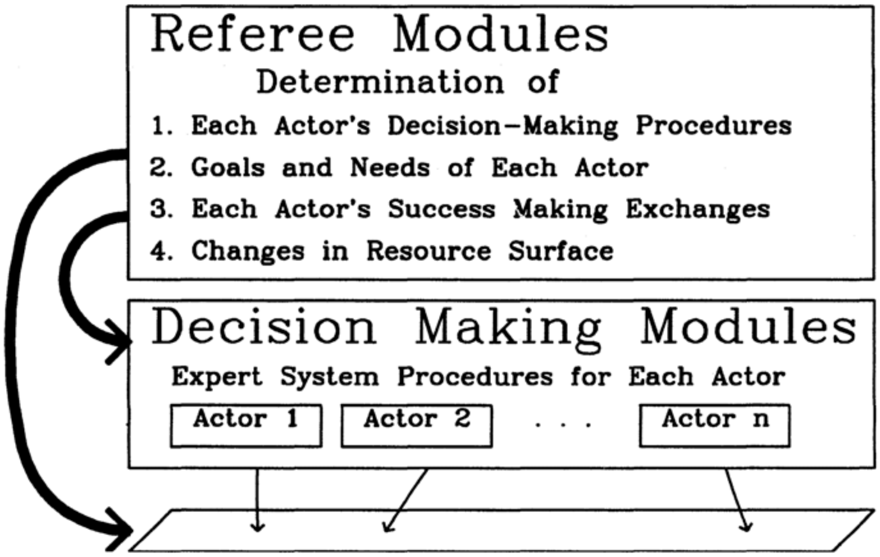


Figure 15.7 Resource surface.

from a predefined library of procedures. Thus, as new 'individuals' are born and mature, one of the referee subprograms will determine exactly which production rules the individual may use. The initial set of production rules will depend upon what the referee determines to be sources of behavioural transmission, but subsequently changes may arise from the influence of behavioural models and from cognitive 'failures'—ie. instances where following the cultural rules leads to an inability to meet the current needs or goals defined for the individual by the referee. These failures may impel the redefinition of the individual's set of production rules even in the face of cultural sanctions.

The specific function of each decision maker expert system is to simulate how an individual will attempt to meet the needs and goals imposed by the referee. The simulated reasoning of the individual will be translated into three kinds of actions—interacting with the resource surface, making exchanges, and seeking information from other individuals about procedures which will allow the individual to obtain desired goods. The specific actions taken will be determined using information supplied by the referee to the individual about the individual's own material and social resources.

The referee modules

The referee subprograms manage the simulation. The use of expert systems in the referee subprograms is tied to the need to make the reasoning used in the simulation management relatively clear and accessible. The referee modules do not need to model accurately the nature of human decision making. There are currently four different major referee subprograms, though this list

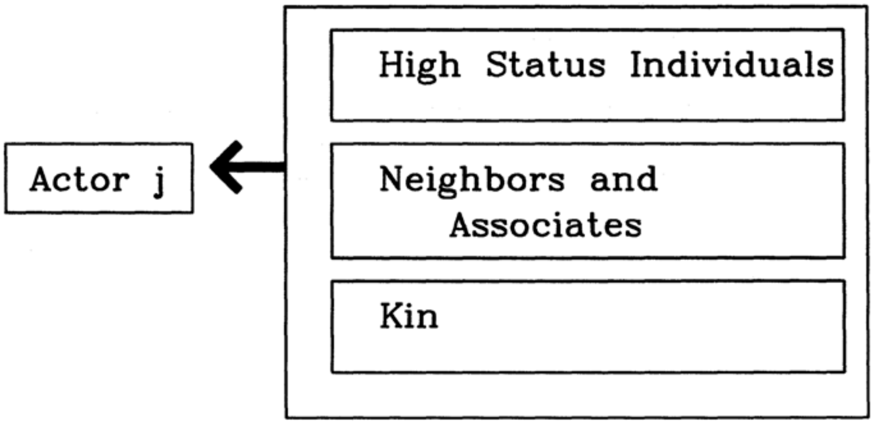


Figure 15.8 Transmission of new ideas.

will probably be augmented in the future. The general role of the referee modules within the whole program structure is given in Figure 15.7.

(1) Procedure Determination: Here the sets of production rules used by individuals will be determined in accordance with the model(s) of the transmission of cultural ideas defined for the simulation run (see Fig. 15.8). Also, modifications in response to events in the simulated world will take place here. One interesting aspect of the construction of this subprogram will be determining how to deal with forces impelling an individual towards the possession of contradictory production rules (see above). The criteria for revising the expert systems of decision makers will be important for modeling the type of situation described by Conrad & Demarest (1984) where existing schemata (in their case, an expansionist ideology) conflict sharply with the apparently rational course of action.

(2) Goals/Needs Determination: The goals and needs of the simulated individuals will depend on a number of factors which are not defined at this time, including the time delay between decision-making events and the role played by the simulated individual within his/her society. The burden imposed on individuals will fall into three general categories: (1) obtaining necessary subsistence goods; (2) obtaining/expanding control over material and social resources; and (3) obtaining prestige for its own sake. In complex societies, the nature of these goals vary qualitatively between different social classes, and so one major problem in developing this subprogram will involve handling the development of these differences.

(3) Exchange Resolution: One of the most difficult parts of the simulation will be resolving the attempts of individuals to exchange goods. Depending on the social context, the availability of exchange partners, and an individual's assessment of the values of the items being exchanged, a party to a potential exchange will accept or reject the exchange. In simpler situations, two decision-makers may provide a sequence of counteroffers depending on their cognized valuation of an item or service in the manner predicted by Edgeworth Box analyses (Lave & March 1975, pp. 203–209). In all contexts, however, a decision-maker may need either to step beyond his/her cognized limits or else accept the consequences of walking away from the exchange. This referee subprogram will need to interact with both potential exchange partners and determine (1) whether the items exchanged are obligatory or freely determined and, if the latter, whether or not the exchange is accomplished through exchanges of gifts, barter, or some other means of exchange; (2) who approaches whom to initiate the exchange; (3) whether or not one or both of the exchange partners will exceed their preconceived notions of cost which subsequently may result in the procedure determining referee revising the expert system representation of one or both individuals.

(4) World Management: This subprogram provides most of the information used by the simulated individuals. The resources available to different individuals are determined here. Different factors, including climatic variability, investment in irrigation works, and resource depletion can be used to determine these resources. This subprogram or a future offshoot will also track who controls different resources, what new resources appear, what acquaintances and potential exchange partners individuals know, who is born, and who dies.

Concluding remarks

Despite the amount of work these kinds of simulations ultimately will require, the potential results justify the expenditure of effort. Archaeologists know very well the limitations on what we can learn directly from the archaeological record about past cultural change. But, instead of ignoring potentially significant phenomena which we cannot observe, we should try systematically to understand the possible implications of various mechanisms of information processing with regard to the kinds of evidence which we can observe archaeologically. Simulations of past cognitive processes and decision making provide a means for improving the theoretical framework from which we generate hypotheses about past cultural change, and few other alternative approaches exist.

References

Boyd, R. & P.Richerson 1985. *Culture and the evolutionary process.* Chicago: University of Chicago Press.

Cavalli-Sforza, L.L. & M.W.Feldman 1981. *Cultural transmission and evolution: a quantitative approach.* New Jersey: Princeton University Press.

Chippindale, C. 1987. Comment on Earle and Preucel 1987. *Current Anthropology* **28,** 513.

Christenson, A.L. 1980. Change in the human niche in response to population growth. In *Modeling change in prehistoric subsistence economies,* Earle, T.K. & Christenson, A.L. (eds), 31–72. New York: Academic Press.

Cicourel, A.V. 1986. Social measurement as the creation of expert systems. In *Metatheory in social science: pluralisms and subjectives,* Fiske, D.W. & Shweder, R.A. (eds), 246–270. Chicago: University of Chicago Press.

Clancey, W. 1983. The epistemology of a rule-based expert system—a framework for explanation. *Artificial Intelligence* **20,** 215–251.

Conrad, G. & A.Demarest 1984. *Religion and empire: the dynamics of Aztec and Inca expansionism.* Cambridge: Cambridge University Press.

D'Andrade, R. 1981. The cultural part of cognition. *Cognitive Science* **5,** 179–195.

D'Andrade, R. 1989. Cultural cognition. In *Foundations of cognitive science,* Posner, M.I. (ed.), 795–830. Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Doran, J. 1982. A computational model of socio-cultural systems and their dynamics. In *Theory and explanation in archaeology,* Renfrew, C., Rowlands, M. & Segraves, B. (eds), 375–388. New York: Academic Press.

Doran, J. 1987. Anthropological archaeology, computational modeling, and expert systems. In *Quantitative research in archaeology: progress and prospects,* Aldenderfer, M. (ed.), 73–88. California: Sage Publications.

Earle, T. 1980. A model of subsistence change. In *Modeling change in prehistoric subsistence economies,* Earle, T.K. & Christenson, A.L. (eds), 1–29. New York: Academic Press.

Earle, T. & R.Preucel 1987. Processual archaeology and the radical critique. *Current Anthropology* **28,** 501–538.

Flannery, K. 1986a. Adaptation, evolution, and archaeological phases: some implications of Reynolds' simulation. In *Guila Naquitz: archaic foraging and early agriculture in Oaxaca, Mexico,* Flannery, K. (ed.), 501–507. New York: Academic Press.

Flannery, K. 1986b. The modeling of foraging strategy: an introduction to part VII. In *Guila Naquitz: archaic foraging and early agriculture in Oaxaca, Mexico,* Flannery, K. (ed.), 435–438. New York: Academic Press.

Hassig, R. 1985. *Trade, tribute, and transportation: the sixteenth-century political economy of the valley of Mexico.* Norman: University of Oklahoma Press.

Hodder, I. 1982a. Sequences of structural change in the Dutch neolithic. In *Symbolic and Structural Archaeology,* Hodder, I. (ed.), 162–177. Cambridge: Cambridge University Press.

Hodder, I. 1982b. Theoretical archaeology: a reactionary view. In *Symbolic and Structural Archaeology,* Hodder, I. (ed.), 1–16. Cambridge: Cambridge University Press.

Hodder, I. 1985. Postprocessual Archaeology. In *Advances in Archaeological Method and Theory,* Schiffer, M. (ed.), **8,** 1–25. New York: Academic Press.

Hodder, I. 1986. *Reading the Past: Current Approaches to Interpretation in Archaeology.* Cambridge: Cambridge University Press.

Johnson, G. 1973. *Local exchange and early state development in southwestern Iran.* Anthropological Paper 51. Ann Arbor: The University of Michigan Museum of Anthropology.

Johnson, G. 1982. Organizational structure and scalar stress. In *Theory and explanation in archaeology: the Southampton conference,* Renfrew, C., Rowlands, M. & Segraves, B. (eds), 389–421. New York: Academic Press.

Johnson, G. 1988. The changing organization of Uruk administration on the Susiana Plain. In *The Archaeology of Western Iran: settlement and society from prehistory to the Islamic conquest,* Hole, F. (ed.), 107–139. Washington: Smithsonian Institution.

Lave, C.A. & J.G.March 1975. *An introduction to models in the social sciences.* New York: Harper & Row.

Lemmon, H. 1986. COMAX: an expert system for cotton crop management. *Science,* **233,** 29–33.

Limp, W.F. & C.Carr 1985. The analysis of decision making: alternative applications in archaeology. In *For concordance in archaeological analysis: bridging data structure, quantitative technique, and theory,* Carr, C. (ed.), 128–172. Kansas City: Westport Publishers.

Miller, D. & C.Tilley (eds) 1984. *Ideology, power and prehistory.* Cambridge: Cambridge University Press.

Miller, G. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* **63,** 81–97.

Politakis, P. & S.Weiss 1984. Using empirical analysis to refine expert system knowledge bases. *Artificial Intelligence* **22,** 23–48.

Read, D.W. 1985. The substance of archaeological analysis and the mold of statistical method: enlightenment out of discordance? In *For concordance in archaeological analysis: bridging data structure quantitative technique, and theory,* Carr, C. (ed.), 45–86. Kansas City: Westport Publishers.

Read, D.W. & C.Behrens 1989. Modeling folk knowledge as expert systems. *Anthropology Quarterly* **62,** 107–120.

Renfrew, C. 1987. Problems in the modelling of socio-cultural systems. *European Journal of Operational Research* **30,** 179–192.

Reynolds, R.G. 1986. An adaptive computer model for the evolution of plant collecting and early agriculture in the eastern Valley of Oaxaca . In *Guila Naquitz: archaic foraging and early agriculture in Oaxaca, Mexico,* Flannery, K. (ed.), 439–500. New York: Academic Press.

Reynolds, R.G. & B.P.Zeigler 1979. A formal mathematical model for the operation of consensus-based hunting-gathering bands. In *Transformations: mathematical approaches to cultural change,* Renfrew, C. & Cooke, K. (eds), 405–418. New York: Academic Press.

Sanders, W.T. & R.S.Santley 1983. A tale of three cities: energetics and urbanization in pre-Hispanic Central Mexico. In *Prehistoric settlement patterns: essays in honor of Gordon Willey,* Vogt, E. & Leventhal, R. (eds), 243–291. Albuquerque: University of New Mexico Press.

Simon, H.A. & C.A.Kaplan 1989. The foundations of cognitive science. In *The Foundations of Cognitive Science,* Posner, M.I. (ed.), 1–47. Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Wright, H.T. & G.A.Johnson 1975. Population, exchange, and early state formation in southwestern Iran. *American Anthropologist* **77,** 267–89.

16

Syntax and semantics of figurative art: a formal approach

COSTIS DALLAS

Introduction

> The relationships which analytical archaeology may hope to elucidate are those found in the separate but interconnected domains of archaeological *syntactics, pragmatics, and semantics*. …One of the more interesting possibilities arises from the chance that one or all of these separate grammars may eventually be expressed in a symbolic and axiomatic calculus. (Clarke 1968, pp. 485–6)

The study of culture as a communication system, functioning in a manner not dissimilar to language, may be classified under the headings of structuralism, structural analysis or semiotics. Three quarters of a century after the publication of Saussure's *Cours de Linguistique Générale,* structural approaches now extend to encompass diverse aspects of human culture, such as the study of kinship, gestures, spatial aspects of human communication, dress, and narrative (Lévi-Strauss 1949; Birdwhistell 1952; Hall 1959; Barthes 1966; Barthes 1967). In linguistics, structural analysis was significantly extended by work on generative grammars (Chomsky 1957; Chomsky 1965), case (Fillmore 1968) and structural semantics (Katz & Fodor 1963; Katz 1971; Lakoff 1971). The semiotics of the visual arts, a field with its own history and current theoretical agenda (Eco 1968; Calabrese 1985), was recently integrated within a general semiotic theory, regarding all cultural processes as processes of communication, each of which presupposes an underlying system of signification (Eco 1976).

The impact of French structuralism on archaeology can be traced back to Leroi-Gourhan's study of the iconography of palaeolithic cave art (Leroi-Gourhan 1965). The early influence of American generative linguistics is also apparent in a number of archaeological studies (Munn 1966; Deetz 1967; Glassie 1975). Current research on the symbolic aspects of material culture regularly expresses archaeological problems in semiotic or structural terms (eg. Hoffmann 1977; Conkey 1982; Taylor 1987; Llamazares 1989), and a structural approach was used to address empirical issues of design systematics (Hodder 1982a; Hardin 1983; Washburn 1983a). Recent critical surveys of the application of structural methods in archaeology can be found in Miller (1982), Hodder (1982b; 1986, pp. 34–54; 1987) and Conkey (1989).

The marriage between structural analysis and formal archaeology has not been altogether a happy one, despite their common intellectual aims. The existence of material culture 'symbols' is often proclaimed in current archaeological research, and 'codes' are invoked to account for their integral organization and for the generation of 'meaning', yet the liberal adoption of structural and semiotic terminology in questions involving the meaning of material culture is rarely matched by operational procedures that would allow the rigorous examination of propositions and theories against empirical data. This frequent lack of formalization creates the suspicion that structural or semiotic explanations often amount to little more than post-hoc accommodative arguments (Renfrew 1982, p. 13), or, still worse, mere exercises in free association.

While the analogy with language, and therefore the applicability of structural analysis, can be more easily defended for figurative art than for other archaeological material (since there should normally be little doubt that representations had a communicative function) the need for the provision of an operational framework, allowing the empirical examination of structural theories, remains equally important. The integration of structural analysis with formal and quantitative archaeology is necessary not only in order to provide the former with dependable means for the discovery and confirmation of propositions, but also in order to provide the latter with a way of dealing with figurative 'texts' which models adequately their formal complexity and recognises explicitly their multilayered, meaning-bearing function.

From the viewpoint of archaeology, structural analysis may, thus, be seen as an approach to data definition, ie. to the formal representation of archaeological information. Clarke, quoted above, called for the definition of operational models ('a symbolic and axiomatic calculus') to represent the underlying systems conditioning the formal configuration ('syntactics'), significance ('semantics') and technological/functional context ('pragmatics') of archaeological entities. The present study approaches Clarke's programmatic goal by introducing a practical framework, intended for formalization using a computer, that should allow the effective investigation of structure in figurative art. The framework takes into account both the complex spatial and topological relationships inherent in visual representations, and the dependence of meaning on a multiplicity of signifying systems operating both on ‘lexical’ elements and on their formal configuration.

A case study of relief representations on Classical Attic grave stelai is used in order to illustrate problems concerning the semiotic formalization of figurative art. The specific aims of the study are:

1. to discuss alternative syntactic approaches to the symbolic representation and classification of figured images.
2. to introduce a set of syntactic relations appropriate for the representation of spatial and topological complexity in figurative art.
3. to identify relevant aspects of semantics and to present a model of semantic interpretation.
4. to suggest a practical approach for the formal treatment of image syntax and semantics using logic programming.

Classical Attic grave stelai

This study presents work in progress on the structure and meaning of figured reliefs on Classical Attic funerary stelai; the stelai, produced in a limited geographic area from around 430 BC to 317 BC, form a well-defined, homogeneous series (Conze 1893; Diepolder 1931; Johansen 1951). With few exceptions, they are now disconnected from their archaeological context; the determination of their function and significance depends, therefore, on the study of their observable traits. Data from several hundred stelai are used for the current project; some typical examples, selected for the illustration of specific points, are presented in outline sketches (Fig. 16.1).

Most sculptured Classical Attic stelai display stereotyped human figures, without individual physiognomic characterization, arranged from left to right in stationary compositions. Inscribed names, corresponding to figures depicted, generally identify individuals belonging to the same nuclear family: husband and wife, sometimes with a child or children; a parent with children; brothers and sisters. Grave reliefs depicting two or more figures have been interpreted in the past as tomb cult scenes, representing the dead in ghostly apparition in front of mourners coming to administer funerary rites in the cemetery; family reunion scenes, representing the deceased ancestors welcoming the newly-dead in the underworld; or farewell scenes, showing the deceased in the company of members of the bereaved family. Irrespective of the specific interpretation of the images, figures on most stelai may be safely identified as ‘idealised representations of the deceased’ and their mourning relatives (Johansen 1951, p. 16).

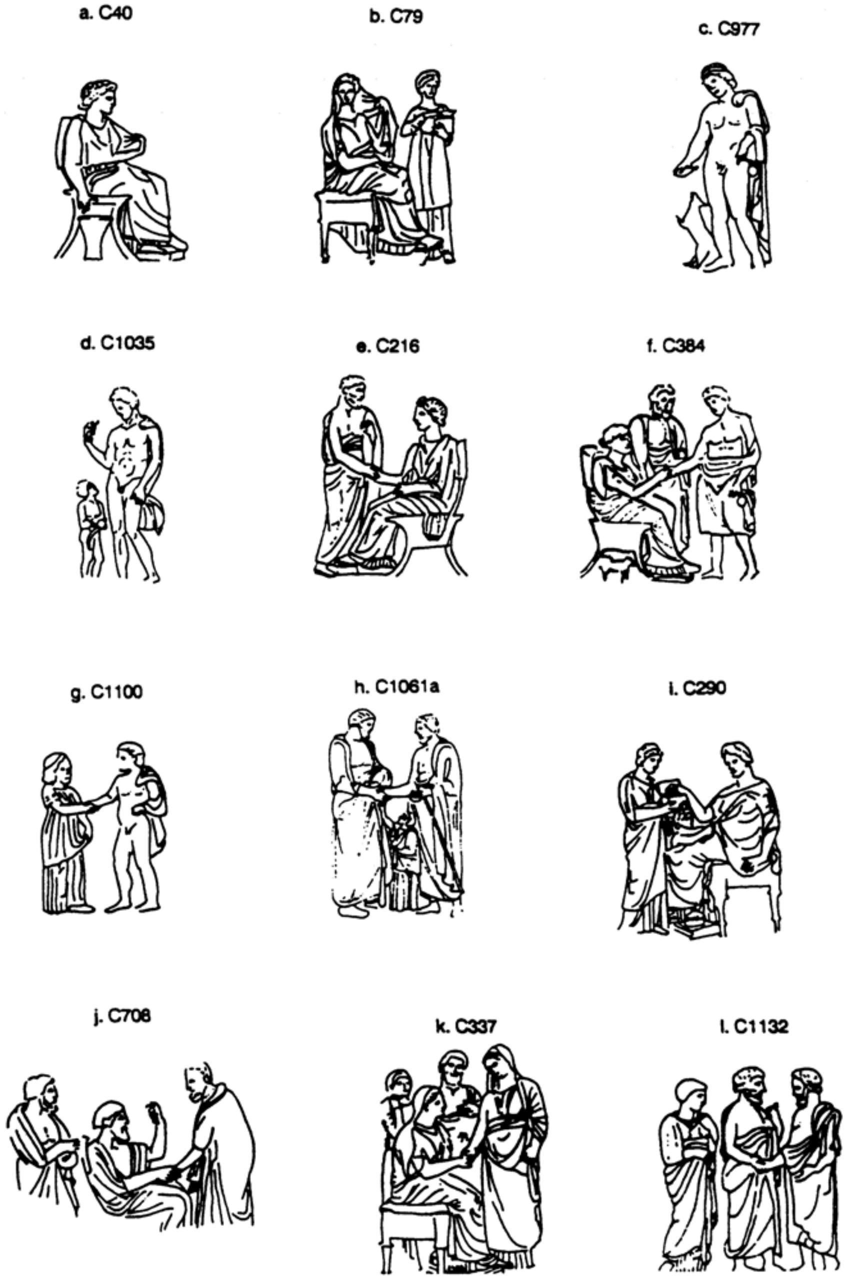
Important recent research focuses on the relationship between the iconography and the conditions of production, use and function of stelai (Clairmont 1970; Schmaltz 1983). Analysis of the social significance of costume on stelai based on the integration of quantitative techniques with role theory (Saxe 1970), sociology (Roach & Eicher 1965) and semiotics of dress and adornment (Barthes 1967) has revealed that the costume types, derived by means of numerical classification, but also the attached *attributa*, posture, gestures and arrangement of figures in the image, were related with socially significant gender/age categories (Dallas 1987). The way human figures were depicted on stelai was also found to be associated with stele expenditure (an association partly due to vertical social differentiation), with ascribed social affiliation, and with the identification of figures (by means of name inscriptions) as the—primary or secondary—deceased.

The morphology of human figures depicted on stelai was, thus, examined as a system of signs, an iconic ‘language’, conditioned by social categorization processes. The *syntactics* of that ‘language’ correspond to an analytically derived typology of figures, based on their observable attributes of dress and adornment, and related to their pose, gestures and attached non-figural elements. Its *semantics* are the composite social identities constituting the social categorization system of Classical Attica. Finally, its *pragmatics* correspond to a historically attested practice of conspicuous funerary expenditure, designed to strengthen family solidarity.

Stele composition

Schefold identified the investigation of the formal language, content and aesthetic function of Classical Attic stelai as a critical factor for their understanding (Schefold 1952). This is not a trivial task. To paraphrase Chang (1967, p. 77), the question is not if the meaning of stelai *can* be inferred from physically observable differences in stele composition. Surely it can; how else could ancient Athenians understand what a particular stele represented? The problem is to provide a descriptive and analytical account of stele composition that will allow the determination of the *functional structure* of images, ie. the internal order that enables them to convey meaning.

The term *composition* has been used in Classical archaeology to represent diverse notions. Apart from its use as a synonym for ‘representation’ or ‘scene’, it has also been employed to denote the geometric arrangement of elementary forms devoid of meaning. For instance, the formal arrangement of the fifth century BC Leukothea relief was defined as a ‘pattern of composition’ composed of ‘a cursive letter ‘W’ in its upper half, with a heavy square set against lighter vertical dashes in the lower’ (Carpenter 1960, p. 328). The principal iconic elements of the stele of Hegeso— heads, shoulders and elbows of figures



a. C40
b. C79
c. C977
d. C1035
e. C216
f. C384
g. C1100
h. C1061a
i. C290
j. C708
k. C337
l. C1132

Figure 16.1 Sketch illustrations of typical Classical Attic grave stelai, identified by catalogue number (after Conze 1900).

—were thought to form a virtual polygon, inscribed in a circle whose centre was the seated matron's hand: the focus of the composition (Dohrn 1957, pp. 107–108). Both examples regard composition as non-figurative pattern; their categories are theoretical constructs, derived from the non-semiotic *stuff* or *continuum* of the sculptural field (Hjelmslev 1943), and are related to a latent geometric structure by direct correspondence, without the mediation of a system of signification.

By extending the figure morphology model presented above, a semiotic approach to stele composition may be formulated. Human figures on stelai—representations of social *personae*—communicate meaning not on their own, but in the context of the entire sculptured image of a funerary marker, annotated by name inscriptions and possibly by an epigram, and erected in front of a funerary family plot. On this account, stele composition constitutes a semiotic system encompassing figure morphology by a process of *overcoding:* human figures, the minimal combinatorial units of stele composition, are the maximal combined chains of a preceding code, figure morphology. Stele composition may thus be defined as the spatial (adjacency-based) and topological (connectivity-based) arrangement of human figures in the image field.

A definition of composition is implicit in the analysis and classification of Classical Attic stelai by archaeologists. The number of figures depicted provides a general criterion of classification (Conze 1893; Schmaltz 1970). Stance, and adjacency relationships between figures introduce a further distinction. Thus, Schmaltz identifies the depiction of a seated figure flanked by two standing figures as a 'Kompositions-Schema' (Schmaltz 1970, p. 53). Connectivity relationships, such as the tactile link between two figures by means of a handshake (Johansen 1951, pp. 54–55), or the visual isolation between the figures (Himmelmann-Wildschütz 1956) are also introduced as compositional categories. The gender and assumed social identity of the figures are included in the compositional description, eg. in the definition of 'mistress and maid' stelai(Johansen 1951, pp. 18–21; Schmaltz 1983, pp. 16–17), and the arrangement of the figures, in combination with an *attributum* such as a trinket box is used to identify a compositional theme, ie. female adornment (Stupperich 1977, pp. 91–94).

Stele composition is seen in Classical archaeology as the juncture of two main axes of variability: style and function. On the first axis, compositional affinity is used for attribution and dating (Diepolder 1931). Stele composition is regarded as an evolutionary process, based on the gradual expansion of simple compositional archetypes (Johansen 1951). The diachronic development of compositional forms involves, however, not only the count of figures in the image, but also their formal interrelationships: visual contact is gradually replaced by visual isolation, opposition by juxtaposition, profile view by frontality and 'opening up' of the image towards the spectator (Diepolder 1931, pp. 21–22, 25; Karouzos 1966, pp. 264–266; Schmaltz 1983, pp. 203–207).

Chronological changes to stelai composition need not necessarily be examined from a formalist perspective. Recalling Riegl's concept of *Kunstwollen,* Himmelmann relates the chronological dependencies of isolation, frontality and nudity or adornment of the main figure with changes in the 'will to art' regarding the conception of the dead and its sculptural embodiment (Himmelmann-Wildschütz 1956). Schmaltz pursues a similar course in his account of the way the semicircular disposition of figures around a central focal point is replaced, in late fourth century BC, by clearly separated foreground and background planes. The change in formal structure is intimately linked with the interpretation of the scene: in the foreground, the 'protagonists' shake hands, while the survivors 'cling' in the background (Schmaltz 1983, p. 206). In effect, compositional structure is used in order to elucidate iconographic content.

The second axis of variability underlying stele composition, meaning, concerns a 'unified treatment of the still unsolved problems of interpretation' (Johansen 1951). A central question of stele iconography is the identification of the deceased. Assuming that stelai depict both the deceased (as they were when still alive) and their mourning relatives, the former should be differentiated from the latter not only by means of figure morphology, but also by means of compositional role. The deceased is thus identified as the seated figure or the figure shown with subordinate elements such as a diminutive servant or a pet dog (Schmaltz 1983, p. 211), or as the figure occupying the centre of the composition, not only on account of spatial arrangement, but also on account of visual contact and gestural direction (Diepolder 1931, pp. 21–22).

Symbolic representations

The identification of adequate symbolic representations of the 'state of the world', amenable to machine manipulation, is recognised as a necessary condition for the development of 'intelligent' systems (Levesque 1986). In archaeology, 'viable' data definition is considered to be of paramount importance for the validity of the operations of comparison, classification and interpretation (Boast 1985, p. 160). Symbolic representation systems have been developed by the French *logiciste* school, among them a componential, purely syntactic code for the description of ornament (Gardin 1978), and semantic codes utilizing a form of *case* grammar for the description of iconographic material (Le Rider 1975; Salomé 1980); recent expert systems studies involving the formalization of archaeological argument concern also the semantics of figurative art (Lagrange & Renaud 1985; Lagrange & Renaud 1988).

If symbolic representations are used in archaeology in order:

> to make way for the retrieval and comparison of morphological data …and…to supply the symbolic elements that go into the formulation of theories (Gardin 1980, pp. 39–40),

then suitability for these tasks may be used as a means of assessing alternative representations. These tasks correspond closely with the following criteria, advanced by artificial intelligence workers for choosing among symbolic representations:

> The representation should provide a way to state any fact in a given domain. It should make intuitively simple facts look simple. It should make intuitively similar facts look similar. Most important, the representation should support useful inferences. (Charniak & McDermott 1985, p. 333)

The four criteria can be summarised as descriptive adequacy, simplicity, analytical adequacy and relevance. So far as stele composition is concerned, a suitable symbolic representation should reflect adequately the way composition is conceived by workers in the field of stele iconography. Symbolic representations of ascending complexity should be examined, starting with the simplest. The representation to be chosen should reflect well the taxonomic relationship between stelai—so that similar images have similar symbolic representations and *vice versa*—and should allow a suitable classification of images. Moreover, it should be possible to express questions of interpretation related to stele composition, such as the definition of the theme represented or the identification of the deceased, in terms implicating the symbolic representation selected, and to provide answers to these questions. By expressing semantic relationships, the symbolic representation of stele composition 'can be regarded indifferently as the representation of an image, or as the representation of a discourse associated with the image: its 'legend', so to speak, in the etymological sense of the word' (Gardin 1980, p. 44).

It should be clear at this point that the analysis of stele composition cannot be confined to the study *in vitro* of its lexical components, ie. of human figures, or, worse, of arbitrarily selected iconographic motifs. To reiterate the findings of *Gestalt* psychology, 'the apprehension and comprehension of visual images is more than the inventory of the parts which form them. The spatial relationships between the parts play a crucial role in the perception of the whole', and are involved in the communication of messages (Lisboa 1985, pp. 187–8). It has been noted that by deleting an element from an image, the meaning of the resulting image cannot be expressed as the difference between the meaning of the original image and the deleted element (Carpettini 1980, p. 99). In analogy with language (Charniak & McDermott 1985, p. 174), the functional structure of images is not explicitly manifested in the enumeration of their parts, but must be *inferred* from their formal configuration.

The required representation should enable us to ask questions not only on the *occurrence* or *co-occurrence* of figures with specific attributes in a given image, but also on their *configuration*, with respect to one another, in the image field. The resulting compositional typology should be useful in explaining:

1. the fact that compositional elements are not combined at random, but present certain regularities, and that some configurations of compositional elements occur exceptionally or not at all.
2. the perceived similarity between different compositional types.
3. the implications of the formal configuration of compositional elements in the image field on the meaning of the image.

The formal configuration of image elements may be regarded as the combination of a structural component, ie. a function of some culturally meaningful latent dimensions of variability, and a contingent component, related to the individual circumstances of particular images (cf. Boudon 1968). The quest for structure can, thus, be seen as a taxonomic enterprise, aiming at the construction of classifications that reflect adequately important dimensions of variability and ignore compositionally unimportant variation; in these terms, the problem is to introduce a symbolic representation of stele composition that will be amenable to such classification.

Stele composition may also be seen as the application of a lexicon upon a categorial component, according to a set of syntactic rules. The lexical component consists of human figure types, corresponding to social *personae* from the Athenian system of social categorization and invested with additional meaning effects related with their elementary *representata*. The categorial component consists of syntactic slots based on attributes such as stance, plane and orientation, and spatial and topological relations such as superimposition and tactile contact, which, when occupied by lexical elements, produce valid instances of stele composition. The determination of stele meaning is an operation of stele syntax on the semantic representation of lexical elements, that is, human figures depicted. A suitable symbolic representation of stele composition may, therefore, be seen as the combination of a suitable representation of the syntactic component with a mechanism of semantic interpretation, ie. the valid amalgamation of individual figure meanings at the image level.

The definition of the syntactic framework of stele composition without recourse to semantics is important, since it provides a criterion for testing alternative syntactic notations, ie. their ability 'to support semantic description;…we shall naturally rate more highly a theory of formal structure that leads to grammars that meet this requirement more fully' (Chomsky 1957, p. 102). By isolating the discussion of formal structure from that of meaning, the syntactic aspect of stele composition can be seen as the context of discovery, its semantics providing the context of validation. This is particularly desirable in archaeological applications of structural analysis, often charged with methodological sloppiness (Hodder 1986, pp. 49–53). Without attempting to answer the vexed issue of the autonomy of syntax (Lakoff 1971, p. 267; Maclay 1971; Weinreich 1971, pp. 314–315), the practical course taken in this study is to attempt a definition of stele syntax independently of, and before the discussion of, stele semantics.

A componential approach

Figures on Attic stelai are arranged in a well-defined left-to-right order. Composition can therefore be visualized as an ordered set of human figure symbols, each occupying a respective position in the image field. Schmaltz (1970, p. 160) used a graphical sketch notation to represent relief images on Classical Attic marble *lekythoi*, funerary vases with similar iconography to stelai. His sketch symbols, intended for efficient visual recall, correspond to combinations of the stance and orientation of full-size human figures.

$$\begin{array}{l} \text{stance} \\ \text{orientation} \end{array} \left[\begin{array}{ccc} s & e & e \\ r & r & l \end{array}\right]$$

$$\begin{array}{l} \text{size} \\ \text{stance} \\ \text{plane} \\ \text{orientation} \end{array} \left[\begin{array}{ccc} f & f & f \\ s & e & e \\ f & b & f \\ r & r & l \end{array}\right]$$

Stele composition may ‘be analysed componentially in terms of distinctive features’, following Hymes (1970, pp. 100–102). Schmaltz’ sketch of compositional types may, thus, be rewritten in a matrix notation, whereby each row represents an attribute or ‘feature’ and each column represents an ordinal left-to-right position in the image. Symbols used for stance are *s* (seated) and *e* (standing); those used for orientation are *l* (left), *r* (right) and *n* (frontal). Using this notation, the image of stele C384 (Conze 1893, no. 384) (Fig. 16. 1f) is represented as:

Figure plane is not explicitly expressed in this matrix, since funerary vases, studied by Schmaltz, as a rule represent all figures in the foreground. Besides, diminutive figures, thought to affect little the visual impression of the image, were altogether omitted by Schmaltz (1970) from his sketch notation.

Images on stelai, on the other hand, are often arranged in a background and a foreground plane. *Plane* is, thus, a necessary part of the matrix representation. For descriptive fullness, half-size figures should also be included in the matrix representation, marked by a further attribute, figure *size*. Since size differentiation on stelai is not a direct reflection of age, but is often an instrument of social—free people *versus* slaves—differentiation (Himmelmann 1971), figure size is best defined at the purely syntactic level. Thus, a four-row matrix, of the form

is necessary in order to represent the composition of C384 (Fig. 16.1f); in this notation, size is marked as *f*(full) or *h*(half), and plane as *f*(foreground) or *b*(background).

Classification

In order to allow classification, stele composition may be recorded in the rectangular data matrix format. Global compositional attributes may be defined on the nominal scale, by concatenating the respective attribute values for individual figures as they appear from left to right. Thus, for the stele configuration represented by the componential matrix above, the global attribute of *size* has the value *fff*, *stance* is *see*, *plane* is *fbf* and *orientation* is *rrl*. Figure count is obviously implicit in these attributes.

The approach to classification adopted here is based on using global attributes, in a preselected order, for partitioning stele configurations. No statistical criterion was used to determine the order of attributes, as in monothetic divisive classification (Gordon 1981, pp. 69–72), since the appropriate order is clearly dictated by their perceived compositional importance. In any case, the order of attributes selected does appear to maximize differences among classes, and the particular choice of classification algorithm does not affect the main argument of this chapter.

A componential classification of stele composition was produced by successively dividing 402 complete grave-relief images on account of figure count—implicit in notations discussed so far—and of the global size, stance, plane and orientation attributes. Although the order of these attributes reflects their relative compositional importance, the results were disappointing. The data were partitioned in 116 classes, representing global size, stance, plane and orientation combinations. No two examples presented in this chapter (Fig. 16.1) were assigned to that same class in this classification. Images with half-size figures were set apart from those without; thus, C290 (Fig. 16.1i) was entirely separated from C216 (Fig. 16.1e). The inclusion of diminutive figures in the data representation both increases the number of types, since, in Classical Attic stelai, these figures adopt an almost arbitrary stance/plane/orientation combination, and obscures the perceived similarity between images.

Rearranging the order of global attributes, so that size accounts for the least general branching operation, has the effect of assigning equal compositional importance to all figures, irrespective of size; for instance, C1035 (Fig. 16.1d) is assigned to a node adjacent with C1100 (Fig. 16.1g), and C1061a (Fig. 16.1h) with C1132 (Fig. 16.1l). However, it is the stance, plane and orientation of *full-size* figures that seems to be intuitively important for the resemblance between images; half-size figures play a secondary formal role in stele composition.

To tackle this problem, figures shorter than 0.7 times the height of the tallest figure in the image were deleted from the symbolic representation; the cutpoint is arbitrary (the ratio is not bimodal), but appears intuitively reasonable. The resulting global compositional attributes of figure count, stance, plane and orientation produce a componential classification in 52 classes; despite the fragmentation, these reflect more adequately the perceived similarity between images. Images with half-size figures are, naturally, classified in the same class as their counterparts depicting only the full-size figures, eg. C1035

(Fig. 16.1d) with C977 (Fig. 16.1c), and C1061a (Fig. 16.1h) with C1100 (Fig. 16.1g). Images with similar left-to-right order of full-size figures— as defined by the stance/plane/orientation combination—are placed together, but mirror images are set widely apart. Thus, C216 (Fig. 16.1e) is put in the same class as C290 (Fig. 16.1i), but both are separated entirely from C79 (Fig. 16.1b), sharing, among global compositional attributes, only the same figure count.

The componential matrix notation, introduced above, is suitable for representing basic formal operations in the analysis of stele composition. Image inversion, ie. the operation producing the mirror image of a scene, can be achieved by reversing the order of the matrix columns, and then changing the sign of the *orientation* symbols, so that *r* becomes *l* and *vice versa*. Inversion may be used to produce a simplified classification that will assign mirror image scenes to single stance/plane/orientation classes. In addition, the number of stele composition nodes can be further reduced by deleting atypical, once-only combinations of compositional attributes from the classification. The resulting classification is presented as a tree diagram (Fig. 16.2), indicating the taxonomic position of illustrated examples.

The diagram shows an intermediate classification structure between a key and a paradigm (Saxe 1970; Whallon 1972). As figure count increases, not all paradigmatic possibilities are realised. In fact, just 12 types, defined by global stance/plane/orientation combinations, account for the great majority (90%) of actual images. Each type corresponds to a variety of images which can be derived by the optional operations of inversion and of half-size figure addition. Excluding this aspect of variability, the taxonomic relationship among images matches well their perceived similarity (Fig. 16.2; cf. Fig. 16.1).

When no branching occurs between successive nodes, attributes thus connected can be said to combine into an ordered complex attribute, since the value of the dominating nodes appear to determine entirely (excluding singleton cases) the value of the dominated node. In fact, orientation is commonly determined by the combination of stance and plane, especially in images confined to the foreground. In practice, therefore, it is possible to ignore figure orientation without loss of the power of the classification in discriminating between significantly different composition forms.

A further elaboration of stele classification, maintaining the purely syntactic treatment of images, consists of deleting all background figures from the compositional representation. The resulting classification contains nine compositional types, defined on the basis of global stance and orientation (Fig. 16.3), and conforms with common judgement on the compositional unimportance of background figures (eg. Schmaltz 1983, p. 21). On this account, C79 (Fig. 16.1b) is classified with C40 (Fig. 16.1a), and all four images showing a standing foreground figure in front of a seated figure are assigned the same compositional class description (Figs 16.1e, 16.1f, 16.1i and 16.1k).

Beyond classification

The omission of half-size figures, mirror images and background figures from compositional representation led to componential classifications of ascending power and compatibility with the accepted view of stele composition. The procedure followed is related to exploratory data analysis. The search for alternative classifications was intended to achieve, through data re-expression, an increase in the amount of compositional information expressed as pattern, and a consequent reduction of residual ‘rough’ that remains unexplained (Clark 1982, pp. 250–251).

If the discovery of latent structure can be related to the separation of the ‘smooth’ from the ‘rough’, it cannot be assumed that the disposition of half-size or background figures, the left-to-right direction of the composition, or those spatial and topological attributes not included in the componential classification introduced above, can be equated to ‘rough’, that is, to unaccounted-for random variation. In fact, the deletion of half-size and background figures amounts to loss of potentially relevant information, since strictly formal regularities are often determined by the presence of such ‘secondary’ figures, or even non-human image elements; for example, outermost figures as a rule face inwards, a fact obscured when (outermost) half-size figures are deleted from compositional representation. Clearly, if componential classifications deemed to be analytically useful (Figs 16.2–16.3) are to pass the test of descriptive adequacy, they must provide a means of relating reduced compositional classes with full compositional descriptions, that is, an adequate account of within-classes variability.

In addition, of the three classifications presented above, (a) including all figures, with mirror images shown separately, (b) excluding half-size figures and singleton configurations, with mirror-images conflated (Fig. 16.2), and (c) excluding half-size and background figures, with mirror-images conflated (Fig. 16.3), none accounts fully for relationships between its own classes, ie. for between-classes variability. In general, the first-order dissection of stelai according to figure count obscures the relationship between images of differing complexity. Besides, important traits of stele composition often run across classes. Through relating full compositional representations with class descriptions, each of the three classifications displays different valid aspects of compositional structure.

In order to achieve a synthesis of the results of separate componential classifications, a direct representation of perceived similarity relationships between images has been attempted. The taxonomic structure of illustrated grave-relief compositions is shown in the form of a labeled relational graph (Fig. 16.4). Example images are represented as nodes, labeled according to the componential matrix notation introduced above. Branches represent perceived similarity relationships between different images.

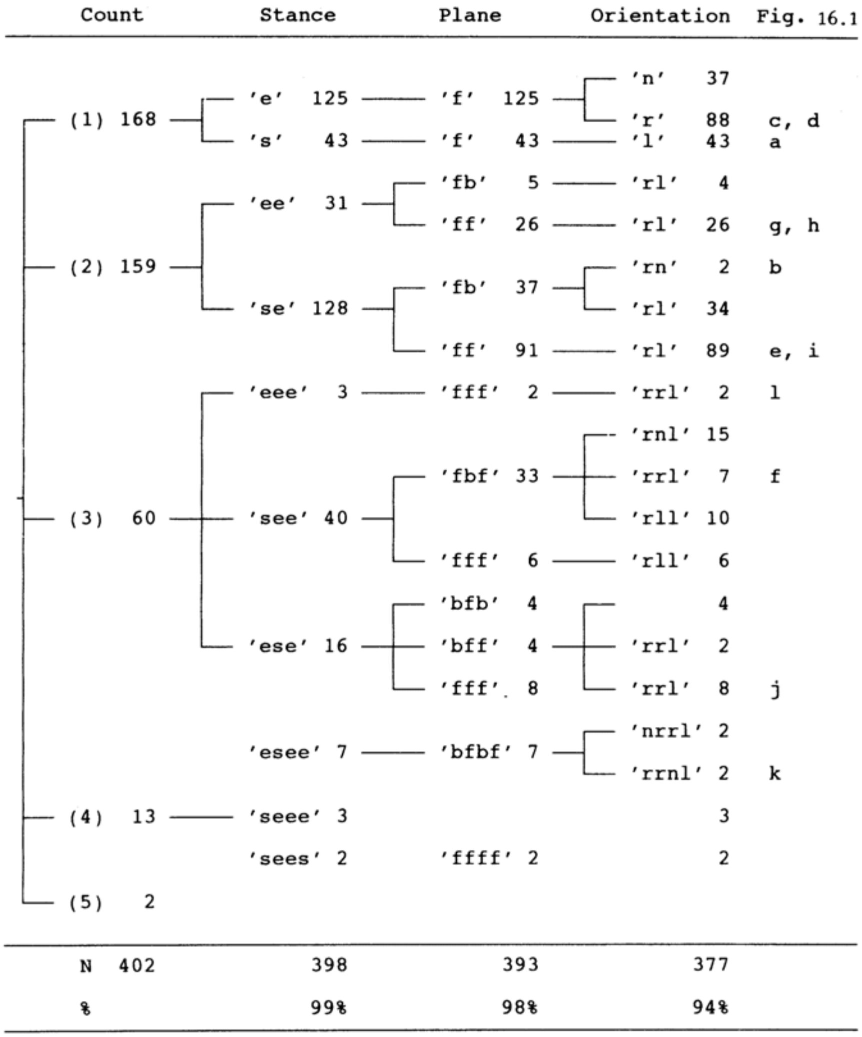


Figure 16.2 Tree diagram of 402 grave-relief images, according to figure count and global configuration of stance, plane and orientation of full-size figures; mirror image nodes are conflated and singletons removed.

The symbolic representation used includes all figures depicted, irrespective of figure size or plane. The graph captures all similarity relationships indicated by componential classifications ignoring half-size figures, image inversion and background figures (Figs 16.2–16.3). Besides, it represents explicitly the effect of half-size figures on compositional similarity, as between C977 (Figs 16.1c, 16.4c) and C1035 (Figs 16.1d, 16.4d); this is important since, despite their secondary role, half-size figures convey compositional meaning (Stupperich 1977, pp. 91–94). In addition, the graph encompasses similarity relationships between increasingly complex compositions, separated in componential classifications on account of the number of full-size foreground figures; for instance, C1132 (Fig. 16.1l, 16.4l) is linked with its compositional subset, C1100 (Fig. 16.1g, 16.4g). The linkage conforms with scholarly opinion, according to which the expansion of images by additional figures 'makes no difference', producing 'a richer version of the old compositional schemes' (Johansen 1951, pp. 17–20, 47). Using just a limited number of examples, this graph provides a fuller account of the taxonomic relationship among images than all three componential classifications taken together.

Transformations

Classifications are *declarative* models, since they are defined by enumeration of their constituent class descriptions, identified by relevant compositional attributes; actual stelai are assigned to a class if their description matches the description of the class. Compositional structure may, however, be more succinctly and powerfully represented if formal relationships are expressed in the form of operations, used as part of a generative mechanism accounting for actual variability. The production of actual images from class descriptions in componential classifications by means of image inversion, the addition of half-size figures (Fig. 16.2) and background figures (Fig. 16.3), and the relationships defined by arrows of the taxonomic graph of

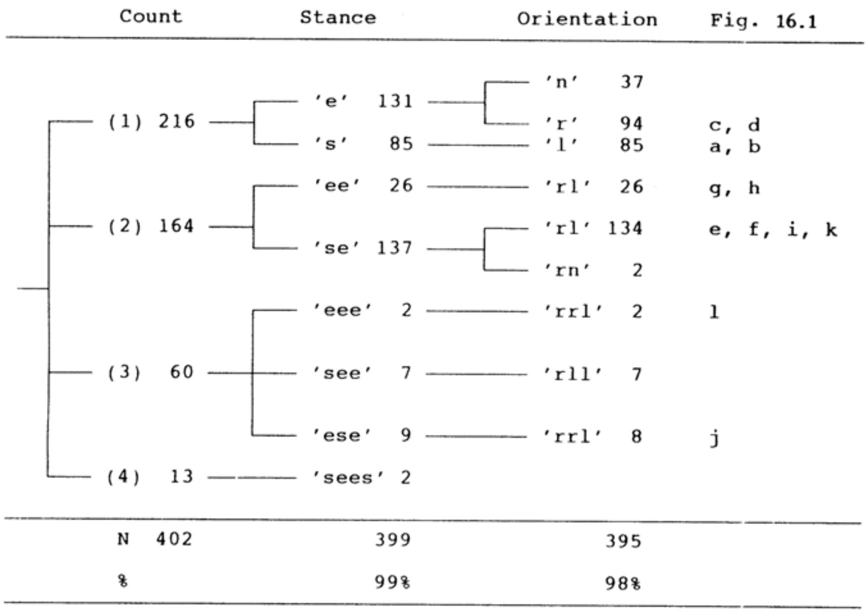


Figure 16.3 Tree diagram of 402 grave-relief images, according to figure count and global configuration of stance and orientation of full-size foreground figures; mirror image nodes are conflated and singletons removed.

$\begin{vmatrix} f \\ s \\ f \\ r \end{vmatrix}$ a. C40 $\begin{vmatrix} f & f \\ s & e \\ f & b \\ r & n \end{vmatrix}$ b. C79 $\begin{vmatrix} f \\ e \\ f \\ l \end{vmatrix}$ c. C977 $\begin{vmatrix} h & f \\ e & e \\ f & f \\ r & l \end{vmatrix}$ d. C1035

$\begin{vmatrix} f & f \\ e & s \\ f & f \\ r & l \end{vmatrix}$ e. C216 $\begin{vmatrix} f & f & f \\ s & e & e \\ f & b & f \\ r & r & l \end{vmatrix}$ f. C384 $\begin{vmatrix} f & f \\ e & e \\ f & f \\ r & l \end{vmatrix}$ g. C1110 $\begin{vmatrix} f & h & f \\ e & e & e \\ f & f & f \\ r & l & l \end{vmatrix}$ h. C1061a

$\begin{vmatrix} f & h & f \\ e & e & s \\ f & b & f \\ r & r & l \end{vmatrix}$ i. C290 $\begin{vmatrix} f & f & f \\ e & s & e \\ f & f & f \\ r & r & l \end{vmatrix}$ j. C708 $\begin{vmatrix} f & f & f & f \\ e & s & e & e \\ b & f & b & f \\ r & r & n & l \end{vmatrix}$ k. C337 $\begin{vmatrix} f & f & f \\ e & e & e \\ f & f & f \\ r & r & l \end{vmatrix}$ l. C1132

Figure 16.4 Compositional similarity between stelai illustrated in Fig. 16.1; images are represented left-to-right, in componential matrix notation (row 1: figure size, row 2: stance, row 3: plane, row 4: orientation).

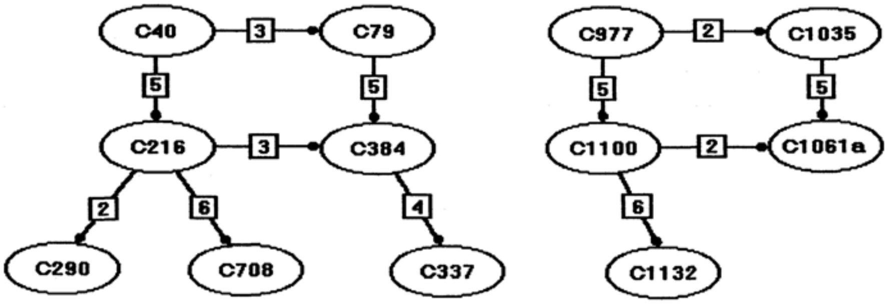
image similarity (Fig. 16.4), imply such a generative mechanism. In this light, classification is used to infer the 'process that generates or recognizes pictures', and becomes a *procedural* model of stele composition (Rosenfeld & Kak 1982, p. 306).

A basic procedural model consists of geometric transformations. The derivation of geometric forms can be achieved by the application of three elementary transformations: translation, rotation and scaling (Newman & Sproull 1981). The relationship between compositional elements may also be expressed as a product of these elementary transformations. Images with more than one compositional element may be seen as concatenated sequences of simpler images; concatenation allows the use of geometric transformations in a generative sense. Thus, the compositional description of C1035 (Fig. 16.1d, 16.4d) is derived from that of C977 (Figs 16.1c, 16.4c) by concatenating the standing figure on the right side of its scaled down, rotated transform. The formalism is attractive because of its apparent simplicity and suitability for operationalization. The number of geometrical operations necessary to produce derivative images from elementary forms could be used as a measure of information content or composition 'cost'; the same measure could be useful in expressing the taxonomic distance between images.

Geometric models are closely related to symmetry analysis (Washburn 1983b). As already noted with regard to that approach (Hodder 1986, pp. 39–40), the apparent objectivity of the method breaks down on closer examination, since transformations depend heavily on description conventions. The expansion of the compositional description by a single attribute, such as the degree of frontality or contact between figures, may produce entirely different sets of transformations for conceptually similar configurations. Besides, small changes to the set of transformations may produce completely different compositional effects, and entirely different transformations may produce similar effects: for instance, using geometric transformations, mirror image variants are generated either by three-dimensional rotation of the image through 180 degrees about its vertical axis of symmetry, or by two-dimensional scaling of the image with a negative argument (Newman & Sproull 1981, pp. 54–56). The transformation space is not isomorphic with the similarity space it generates. It cannot, therefore, be used to describe unequivocally relationships between images, unless the representation is extended by conversion operations between related geometric representations (cf. Charniak & McDermott 1985, pp. 155–156), a fact negating the apparent simplicity of the approach.

Geometric transformations are concerned with formal pattern on the image plane. However, similarity relationships presented above (Fig. 16.4) suggest that sculptured grave stelai are not regarded as mere formal arrangements of image parts, but as representations of *scenes*, three-dimensional configurations of real objects. Therefore, in order to define a syntactic component appropriate for the analysis of meaning, stele configurations of compositional elements should be seen as transformations of scenes. To support inferences about their meaning, these objects—representations of social stereotypes—should be characterized and related by both spatial and physical properties (Rosenfeld & Kak 1982, pp. 316–317).

The graph of perceived similarity relationships (Fig. 16.4) may be re-expressed as a transition network, representing the state space of compositional configurations related by transformation operations (Fig. 16.5a). In representing the taxonomic structure of illustrated examples as a transition network, nodes are now connected by arrows instead of branches; each arrow is labeled by the binary operation necessary for the transition from a compositional node (ie. symbolic representation) to another. Arrow direction implies increasing compositional complexity between images. Five operations, or transformation rules, are defined: image inversion, and variants of the insertion of an additional human figure relative to a base figure



(a)

Rule	Transformation description
1	Image inversion
2	Half-size standing figure addition, in front
3	Background standing figure addition, in front
4	Background standing figure addition, behind
5	Foreground standing figure addition, in front
6	Foreground standing figure addition, behind

(b)

Base	Stele description
a	Seated figure (facing right)
c	Standing figure (facing right)

(c)

Figure 16.5 (a) State transition network of stele configurations illustrated in Figure 16.1; (b) Compositional transformation rules denoted by arrow labels; (c) Base configurations.

(Fig. 16.5b). The insertion operation does not presuppose strict adjacency; for example, the addition of a foreground figure in front of the base is intuitively recognised in actual images irrespective of the possible intervention of a background figure. All nodes representing illustrated stele examples (Fig. 16.1) may be generated by applying subsets of the five transformation rules on two elementary compositional configurations (Fig. 16.5c). The transition network accounts well for the evolutionary view of stele composition, according to which complex, multi-figured configurations are produced by the expansion of simple scenes (Johansen 1951).

The state space representation (Fig. 16.5a) suggests that the generative system of stele composition can be operationally defined as a non-deterministic finite automaton (Bratko 1986, p. 99). Its initial state consists of an empty scene. Input strings are concatenations of symbols representing the selection of a base configuration, image inversion and figure addition operations. Final states are empirically attested compositional configurations, reached when a valid sequence of transformation operations is applied on the initial state. For example, using the symbols introduced above (Figs 16.5b, 16.5c), the transformation string ‘a51’, using the base configuration *a* (seated figure, facing right) and transformations *5* (foreground standing figure addition in front of base) and *1* (image inversion), generates the configuration of stele C216 (Figs 16.1e, 16.4e), and ‘c251’ generates C1061a (Figs 16.1h, 16.4h).

These input strings are equivalent to symbolic representations of stelai shown as final states of the transition network. They, therefore, also constitute a declarative model, partitioning compositional description into a basic core component and an adjunct (Munn 1966). Apart from image inversion and the definition of the base component, denoted by global compositional attributes, other transformation operations correspond to structural slots in the scene, defined in spatial relation to the base. Instead of regarding stele composition as an expansion of a minimal configuration, the state space may be read in the inverse direction, so that actual images are reduced combinations of potentially realisable structural slots. Structural slots can be rewritten as a global presence/absence attribute list, and be used in conjunction with base configuration and image inversion to determine the compositional schema of illustrated examples (Fig. 16.6).

Composition	*Base*	*Inversion*	*Figure addition operations*				
		1	*2*	*3*	*4*	*5*	*6*
C40 (Fig. 1a, 4a)	a	–	–	–	–	–	–
C79 (Fig. 1b, 4b)	a	–	–	+	–	–	–
C977 (Fig. 1c, 4c)	c	+	–	–	–	–	–
C1035 (Fig. 1d, 4d)	c	+	+	–	–	–	–
C216 (Fig. 1e, 4e)	a	+	–	–	–	+	–
C384 (Fig. 1f, 4f)	a	–	–	+	–	+	–
C1100 (Fig. 1g, 4g)	c	+	–	–	–	+	–
C1061a (Fig. 1h, 4h)	c	–	+	–	–	+	–
C290 (Fig. 1i, 4i)	a	+	+	–	–	+	–
C708 (Fig. 1j, 4j)	a	–	–	–	–	+	+
C337 (Fig. 1k, 4k)	a	–	–	+	+	+	–
C1132 (Fig. 1l, 4l)	c	–	–	–	–	+	+

Figure 16.6 Compositional schema of examples illustrated in Figure 16.1, according to a core-adjunct model; symbols used are listed in Figs 16.5b–c.

Syntactic description

In a recent study, Doran expressed doubts about the suitability of the strongly typed, rectangular data matrix for the description of complex archaeological entities:

> There is something deeply implausible in the idea that, say, a bronze dagger or a hunting encampment can be represented as a set of attribute values even if we are prepared to use a full range of value scales. Both a dagger and an encampment involve a nested structure of parts, relationships and functions. Although their description may require a set or array of values, this cannot be the whole story. (Doran 1986, p. 28)

The main problem which figurative art shares with these entities is that figured scenes do not display a fixed internal structure. Unlike individual figures, entire scenes on stelai consist of a variable number of elements, related to one another in a complex manner not amenable to formalization using the traditional case-attribute-value framework. Global attributes used in componential classification (Figs 16.2–16.3), but also in the taxonomic representation involving structural slots (Figs 16.4–16.6), are clearly derivative, involving the sequential or contextual description of individual figures.

A suitable symbolic representation for descriptive purposes should allow scenes to be interpreted in terms of models such as the componential classifications, the syntactic slot declarative model and also procedural models, such as grammars. This representation should not, however, require loss of information or the *a priori* adoption of a specific theory on the structure of stele composition. On the contrary, it should involve the explicit description of all compositional elements and their relationships within a scene.

It has been noted that pictures ‘can generally be represented by a graph structure in which the parts correspond to nodes, labeled with lists of property values…and the arcs are labeled with lists of relationship values’ (Rosenfeld & Kak 1982, pp. 304– 305). Besides, two-dimensional strings may be used for the symbolic description of pictures, providing for efficient matching operations without significant loss of spatial information (Constantopoulos, Drakopoulos & Yeorgaroudakis 1990, p. 336). Since compositional elements of Classical Attic stelai are arranged in a definite left-to-right order, atomic traits of compositional elements and left-to-right precedence relationships betweeen elements can be adequately represented by a labeled, one-dimensional string. The compositional matrix formalism used above is, in fact, a mere notational variant for a string.

In adopting a string representation, nodes may be initially labeled by a four-place predicate, with arguments denoting figure size, stance, plane and orientation, using symbols introduced above. Further compositional attributes, pertaining to single figures, may also be added. The degree of *frontality* of figures, ranging from profile to fully frontal, can be used to qualify figure orientation. Besides, the *head position,* the *gestures* of both arms and the *legs position,* introducing finer distinctions to the general determination of pose, may also be recorded. The scope of these attributes is the single compositional element, and they may, therefore, be included as further arguments of the predicate used to label the nodes of the string representation.

The spatial arrangement of figures on stelai determines a set of relationships concerning spacing and superimposition. The *spacing* of two figures is defined as the axial distance between them, scaled by an appropriate measure such as the maximum figure height. Since figures are strictly left-to-right ordered, a full description of this aspect of spatial organization is achieved by recording the distance of adjacent pairs of figures; ie. for an n-figured scene, only *n*–1 spacing relations need be recorded. Spacing can be easily integrated within the string symbolic representation as an attribute of arrows connecting adjacent figure nodes.

Superimposition, on the other hand, concerns the extent of overlapping, or occlusion, between figures. It is also a binary relation, and is determined by the extent of overlap between the enclosing rectangles of a pair of figures; it may be defined as a numeric measure, or as a reduced nominal index. Unlike spacing, superimposition is an asymmetric relation, applicable to an ordered pair of figures: while the spacing of figures *(a, b)* is equal to that of *(b, a),* their superimposition relations have opposite signs. Besides, it is obvious from examples such as C337 (Fig. 16.1k) that non-adjacent figures may also enter in superimposition relations. For these reasons, the string representation incorporating left-to-right sequence information is here not adequate. Since it is not guaranteed that every figure of a grave-relief overlaps with at least another figure, the description of superimposition relations for a given image will consist of one or more subgraphs of a tree, whose nodes are the figures depicted.

Unlike spacing and superimposition, operating on a two-dimensional projection of the image, *contact* between pairs of figures is based on a conception of images as actual scenes. While not strictly syntactic, the formal recognition of the contact relationship is based on universal perception, rather than on a conventional system of signification. In a common case (eg. C384: Fig. 16.1f), a figure is looking at another who, at the same time, shakes hands with a third one. Two forms of contact are recognised in this example: *visual* contact and *tactile* contact. *Mediated contact* exists when figures are connected through an object linked with both (eg. C290: Fig. 16.1i). The handshake is an example of symmetric contact, but, in the general case, contact is an asymmetric relation. Contact relationships in a scene may be represented by subgraphs of a multigraph, ie. a graph whose nodes are the set of compositional elements depicted, where more than one directed branch (arrow) is allowed between each pair of nodes (Rosenfeld & Kak 1982, p. 313).

Categories introduced so far are distinguished into compositional attributes and compositional relations; the scope of the former is an individual figure, while that of the latter is a pair of figures. Both figure attributes and binary relations are primary data, and should therefore be part of the descriptive symbolic representation for stele composition. An appropriate data structure is a list, the members of which represent left-to-right ordered compositional elements; each compositional element is described by atomic attributes (eg. stance), sets of atomic attributes (eg. gestures) and, to express compositional relations, sets of two-place predicates, identifying, apart from the relationship, which is the related figure.

Syntactic analysis

The symbolic representation of image description, presented above, is suitable for further syntactic analysis. Inputs to the analysis are symbolic descriptions of stele composition; the mechanism of analysis consists of procedures that process the compositional description in order to produce derivative propositions about the data. Thus, while syntactic description consists of primary data, syntactic analysis functions are rules producing derivative data.

Since syntactic description takes the form of a list of left-to-right ordered compositional elements, consisting of atomic attributes and sets of binary relations, analytical functions could be devised that operate on these data structures. The following vocabulary of analytical functions is based on a modification of Leroi-Gourhan’s set of spatial categories, used for the analysis of palaeolithic cave art (Leroi-Gourhan 1982, pp. 19–37). Most may be used either in a query mode, returning a set of matching images, compositional segments or elements from a data base, or in a production mode, generating derivative configurations of compositional elements.

Between-scenes functions

Congruence is the fundamental retrieval function for stele composition, expressing the matching operation of two compositional segments over a set of compositional attributes and relations. Compositional segments range from entire images to single elements; the congruence set ranges from the basic attributes of figure size, stance, plane and orientation used above for classification, to an arbitrary combination of compositional attributes and relations. Minimally, *attribute congruence* can be defined on the basis of matching over a single compositional attribute or relation; for instance, the scenes of C708 (Fig. 16.1j) and C1132 (Fig. 16.1l) have the same set of contact relations, despite their other compositional differences.

Basic image congruence (BIC) indicates that two full images have the same compositional description according to figure size, stance, plane and orientation; enhanced with matching over the set of contact relations between pairs of figures, it provides an effective means of identifying images with the same gross configuration of compositional elements. *Full-size BIC* is defined as the basic image congruence of reduced symbolic representations, after the deletion of half-size figures (eg. C1100: Fig. 16.1g, and C1061a: Fig. 16.1h); *foreground BIC* is the same operation over symbolic representations where all background figures have been deleted (eg. C384: Fig. 16.1f, and C337: Fig. 16.1k).

Figure congruence, on the other hand, is defined as the matching of single figures, in the same image or across different images, over either their size, stance, plane and orientation *(basic figure congruence)*, additionally the figures' compositional relations, or an arbitrary set of attributes and relations. Especially when the congruence set is extensive, this function allows the identification of important formal similarities, ignoring compositional context (eg. C977: Fig. 16.1c, and C384: Fig. 16.1f, rightmost figure). *Segment congruence* is defined if compositional segments—sequences of more than one compositional element—match over compositional attributes or relations. This form of congruence is used to identify scenes that are linear expansions of more basic configurations, such as C1132 (Fig. 16.1l), whose segment excluding the leftmost figure reproduces the basic compositional attribute and contact relationships of C1100 (Fig. 16.1g).

Context congruence is defined as the congruence of the left and right compositional context of two segments, over a given set of attributes or relations. The context is either local (involving the directly adjacent elements in a syntactic representation) or global (involving all remaining figures). An example of local context congruence over the basic compositional attributes and contact relations is shown by the background male figures of C384 (Fig. 16.1f) and C337 (Fig. 16.1k) respectively; full context congruence over the same set is shown by the central figures of C708 (Fig. 16.1j) and C1132 (Fig. 16.1l).

Image inversion, defined for the needs of componential classification, may be extended to operate on the full symbolic representation of composition. *Mirror reflection* is the syntactic analysis function defined by the inversion of figures, compositional segments, or, typically for the needs of our analysis, full images. Combined with different variants of congruence, mirror reflection provides a powerful mechanism of identifying affinities between mirror stele configurations, involving a variable set of compositional attributes and relations.

While between-scenes functions presented above are based on the syntactic representation of stelai on account of left-to-right adjacency, an important class of syntactic analysis functions is defined according to structural *homology* between figures, as is shown by their structural slot relationship with the base figure (Fig. 16.5b, 16.5c), or the contact relations in which they participate. Figure homology is distinguished into local and global, in the same way as context congruence above. For instance, the rightmost figures of all three scenes of C384 (Fig. 16.1f), C1061a (Fig. 16.1h) and C708 (Fig. 16.1j) are locally homologous with respect to the base figure, but only the first and the last are globally homologous, if their contact relation with the remaining third figure is taken into account. Loosely speaking, homology within a local context is exhibited by the seated figure of C290 (Fig. 16.1i) and the figure on the right of C1035 (Fig. 16.1d), both being the base figure, in contact with a half-size figure. Global *image homology* is equivalent to the exact matching of the contact multigraphs of the two scenes, so that the base figures match. Although it does not take into account specific compositional attributes, homology is an important tool of syntactic analysis, highlighting the topological, rather than spatial, structure of stele composition.

Within-scenes and summary functions

Image symmetry is identified when an image can be divided into two segments, such that one is the mirror image of the other. These segments will most often be single compositional elements, such as the standing figures of C1100 (Fig. 16.1g). Symmetry may also be defined as an attribute of a compositional *segment*, ignoring part of the scene. Like congruence, symmetry is defined as an operation over a set of compositional attributes within the same image; for instance, the standing and seated females of C109 are in *attribute symmetry*, on account of head orientation, arm position and the gesture of lifting the mantle at shoulder height with opposite hands. Symmetry is, therefore, to parts of a single image what mirror reflection is to two different scenes. It should be noted that symmetry is here defined at the syntactic level, ignoring subject-matter; it is therefore akin to Leroi-Gourhan's *mass symmetry*, rather than *mirror symmetry* (Leroi-Gourhan 1982, p. 25). *Parallelism*, on the other hand, is the complementary operation to symmetry, defined as the division of the image in two segments, such that one is congruent to the other. All qualifications applying to symmetry are also relevant to parallelism.

Juxtaposition and *opposition*, on the other hand, may be regarded as variants of symmetry and parallelism respectively, defined over the orientation and degree of frontality of the head and body of constituent figures. Figures facing each other, ie. symmetric with respect to orientation and frontality, are said to be in opposition (eg. C1100: Fig. 16.1g). Frontal figures, ie. parallel with respect to these attributes, are in juxtaposition; the figures of C79 (Fig. 16.1b) can be said to be quasi-juxtaposed, if the body orientation of the seated woman is ignored. Opposition implies visual contact, while juxtaposition implies isolation (Diepolder 1931, pp. 22–22). While these notions are syntactic, they have implications for iconographic interpretation (Himmelmann-Wildschütz 1956).

Since its elements are left-to-right ordered, an image can be seen as a vector (Eco 1976, p. 240). *Image directionality* is the absolute, left-to-right spatial direction of the scene, related to an implied temporal order (Conkey 1982, pp. 122– 123). Directionality explains why mirror image scenes are not equiprobable. It is operationally identified on account of the orientation of the base figure, as defined above in relation to the taxonomic view of perceived similarity relationships between scenes. For example, C79 (Fig. 16.1b) and C708 (Fig. 16.1j) have right directionality, while C290 (Fig. 16.1i) and C1035 (Fig. 16.1d) have left directionality.

Directionality has been the subject of symbolic and other interpretations. It is related to the left:right opposition, an important symbolic dimension in ancient Greek thought, homologous with other fundamental binary oppositions (Lloyd 1962). In Greek art, being on the proper right of the central figure—that is, in the left of the image as seen by the spectator—is a sign of good omen, seniority or assumed importance (Ashmole 1972, p. 29). On the other hand, it has been claimed recently that scenes in Greek art must be 'read' in the direction of Greek writing, that is, left-to-right (Small 1987). The most common form of contact between figures on Classical Attic stelai is *dexiosis*, shaking the right hand (Davies 1985). The study of directionality in stele composition reveals interesting aspects of the interpretation of the scenes, which are, however, beyond the scope of the current study.

Finally, global image attributes are derived by summary functions, operating on descriptive syntactic attributes and relations. *Image density* may be defined as a quantitative summary of the spacing between adjacent compositional elements; alternatively, density is defined directly at the image level, as a function of maximum figure height, image width and figure count. On the other hand, the connectedness of the multigraph representing contact relationships between figures in a scene may be used as a summary measure of *scene cohesion*, to differentiate images such as C290 (Fig. 16.1i) from those like C79 (Fig. 16.1b).

Semantics

The field of image semiotics is very complex, with important issues still the subject of heated debate. Open issues of image semiotics concern the levels of articulation, and the relationship of images with verbal language. Besides, the Peircean tripartite distinction between index, icon and symbol, common in archaeological semiotics (eg. Hodder 1986; Taylor 1987) has been in the centre of theoretical controversy (Eco 1976, pp. 178, 191–231; Carpettini 1980, pp. 100–101; Calabrese 1985, pp. 120– 140). Although it can be hardly denied that meaning is communicated by images on account of analogical congruence or similarity, it is difficult to find any non-trivial aspect of pictorial meaning that does not depend, at least in part, on a conventional system of signification. Images are able to generate a variety of meaning effects, some of which may be indexical, others symbolic, others iconic. Rather than contribute to these theoretical questions, the aim of the following discussion is to provide an operational framework for further examination of the issues concerning figurative art semantics, in the backcloth of our practical project.

The reason of establishing a syntactic representation for stele composition was to provide a basis for examining how compositional configurations communicate meaning. Apart from its place in syntactic representation, each scene should, therefore, be assigned a semantic interpretation. For this purpose, it is necessary to define the sign-vehicles of stele composition, that is, the formal expressions that are the carriers of meaning, to identify pertinent semantic fields, ie. the interrelated, ` culturally conditioned sub-systems to which meaning effects belong, and to determine an appropriate symbolic representation of semantic units and of their articulation with syntactic information.

As with the determination of syntactic representation, this discussion is confined to Classical Attic stelai. This does not imply the non-existence of a broader semantic system, covering all relief sculpture or indeed all Classical figurative art. In fact, the semantics of Classical relief sculpture or figurative art can be defined from the semantics of stelai by relaxing constraints of well-formedness valid for actual images as the empirical domain is gradually broadened. The determination of broader models relies less on domain-specific knowledge, and is easier than that of more specific models (Rosenfeld & Kak 1982, p. 21). However, limiting discussion to stelai has considerable advantages. Since they were monuments with a well-defined pragmatic function, their semantics can be determined more precisely than that of wider, functionally heterogeneous assemblages of figurative art, and a more rigorous examination of the efficacy of the semantic system in communicating meaning is possible.

Sign-functions

A first class of sign-functions in stele composition consists of figure attributes, that is, a selection set of values, linked by a paradigmatic, rather than syntagmatic, relationship. Although they are defined according to morphological description conventions, these attributes identify socially relevant aspects of the human figure, such as dress and adornment, gestures, and associated objects. Syntactic attributes may also bear meaning; for instance, the seated stance carries a message of seniority or respectfulness. The semiotic function of figure attributes consists of a juncture of plastic forms, morphologically identified by archaeologists, with their *emic* significance.

A second class of sign-functions consists of human figures, ie. representations of composite social identities (Dallas 1987). Physiognomic and anatomical attributes may identify a specific gender and age; this identification is made possible by a universal perception mechanism (cf. Washburn 1983b, p. 2), based on the realistic depiction of the human figure in Classical art. Dress and adornment attributes communicate the composite social identity of individuals corresponding to each figure, as it was conceived by their family at the time of death. Other attributes are also associated with gender, socially constituted age categories, rank and other social affiliations. Although particular attributes need not participate in the generation of all meaning effects, the social identity of each figure is, in general, signified by the configuration of attributes at the figure level.

Relations between figures, which appear in syntactic description or are derived by means of syntactic analysis, constitute a third class of sign-functions. Superimposition, for instance, implies relative importance between figures, while contact relations between figures signify actions in which they are implicated. On the other hand, the presence of figure attributes such as gestures is sometimes dependent on the immediate or image-wide syntactic context, and may be articulated in analytical syntactic functions, such as symmetry. In such cases, syntactic relationships participate in the generation of meaning, eg. when a figure's gesture of lifting the mantle at shoulder-height is mirrored by a symmetric gesture of a complementary figure.

All sign-functions produce semantic units contributing to the global meaning of the image. A distinction may, however, be drawn according to the form of semantic unit produced by each type of sign-function. Semantic units pertaining to attributes or individual figures are atomic concepts or descriptions, corresponding to words of verbal language (eg. 'strigil', 'seated man'). On the other hand, semantic units based on compositional relationships may also be propositions, in the same way as verbal sentences ('*x* shakes hands with *y*'). Configurations of figures linked by means of syntactic relations and analytical functions, should, therefore, be identified as *super-signs,* which are, in their turn, 'susceptible of further combination in order to produce more complex texts', that is, stele images (Eco 1976, pp. 231–232).

Fields of connotation

The denotative field of figure semantics has been identified as composite social identity. Determination of the identity of a figure involves, however, consideration of the set of both denotative and connotative meanings, produced by individual figure attributes.

The distinction between denotation and connotation is not related to a supposed emotional function or vagueness of the latter, but on the fact that it relies upon the pre-existence of the former (Eco 1976, pp. 54–56). The denotation of an attribute such as handshake is made possible by a graphic code establishing a sign-function, ie. a correlation between a sculptural form (expression) and the concept of 'handshake' (content); its connotation is also a sign-function, whose expression is the previously established sign-function 'handshake', and its content is the concept of 'family solidarity'. Connotation is constituted by the connotative codes which establish it: not only semiotic codes, but also 'secondary modelling systems', social conventions and symbolic systems operating in a given historical context.

A major field of connotation is related to a social connotative code. It is established by the correlation of objects, gestural and proxemic behaviour, as witnessed in stele composition, with particular settings in Classical Athenian social practice. For example, a strigil signifies an activity complementary to physical training; reaching for a trinket box signifies female adornment. These attributes signify also the location in which the activities typically take place, the *palaistra* and the *gynaikonitis* respectively. The connotation of activities or settings is central to the identification of social identity: through those, the strigil of C977 (Fig. 16.1c) identifies the figure as an athlete, a juvenile of *palaistra* age, and the gesture and *pyxis* trinket-box of C290 (Fig. 16.1i) identify the seated figure as an Athenian matron. Additional second-order connotations, enabled by the social function of figure attributes, refer to models of adornment or modesty and the expression of mood.

Figures in Classical Attic stelai may be regarded as adaptations, in pose and costume, of types from major Classical statuary; this view conforms to a long-standing paradigm for the evolutionary study of Greek art (Lippold 1923). Thus, the juvenile male depicted in C1100 (Fig. 16.1g) belongs to a class of figures shown carrying a bunched mantle in the same way as Classical Greek statues of the Richelieu (Arnold 1969, p. 186, pl. 13c) and Andros-Farnese Hermes types (Karouzou 1969). The repetition of the posture and costume type is so close that one could talk of a *citation* (cf. Strocka 1979, p. 163). If Classical figurative art is seen as a corpus of 'texts', this type of connotation, whose expression and content are both visual sign-functions, is enabled by a mechanism of *intertextuality*.

Another connotative field of stele iconography, extensively studied in Classical archaeology, is Greek religion and mythology. Instead of adopting a formal evolutionary interpretation of the intertextual reference of C1100 (Fig. 16.1g) to the sculptural types mentioned above, the standing male with the bunched himation may be taken to connote the god Hermes. This association produces a second-order connotation, that of death, since Hermes, in his incarnations as Chthonios and Psychopompos, is related with the Underworld. The validity of such attributions of meaning, *de rigueur* in Classical archaeology (eg. Thimme 1964; Karouzos 1966), depends, of course, on the practical relevance assigned to a religious-mythological connotative code in the domain of funerary monuments.

Information content (Wobst 1977) can also be seen as a connotative meaning effect of figured images. As noted above, Classical Attic stelai belonged to a system of conspicuous expenditure. That system did not function only directly, through the cost of the grave marker, but also indirectly, through the complexity of the image, which may thus be equated to symbolic 'cost'. The complexity and adornment value of costume, in particular, has been identified as a major vehicle for the advertisement of social identity (Roach & Eicher 1965; Bogatyrev 1938).

It is important to note the existence of inversions in the connotation of conspicuous expenditure. In Classical Attic stele composition, inversions may be due to the operation of an egalitarian ideology that tends to obscure social differentiation (cf. Miller 1982), or to an ideology of modesty. It may also be due to the operation of further levels of connotation. A case in point is nudity, which, apart from its direct signification of athletic identity, may connote the heroic ideal (Himmelmann-Wildschütz 1956; *contra* Clairmont 1970); nude heroes are often shown in Classical Greek art in battle with dressed opponents, and there are occasional parallels of the theme in funerary art. A very simple form, low in morphological information content, may hide a complex web of connotative meanings.

According to a classic distinction (Jacobson 1956), connotation is enabled by two fundamental mechanisms of signification: metonymy or metaphor. Metonymy operates on the (syntagmatic) axis of combination, while metaphor operates on the (paradigmatic) axis of selection. Metaphor is illustrated by C1100 (Fig. 16.1g): the male with the bunched himation stands metaphorically for Hermes; similarly, the seated woman of C79 (Fig. 16.1b), unveiling herself by a gesture associated in Classical art with marriage, is a metaphor for a bride, or even for the goddess Hera (Karouzos 1966). Metonymy, on the other hand, is illustrated by the trinket box depicted in C79 (Fig. 16.1b) and C290 (Fig. 16.1i), standing for the act of female adornment, and by the *pais* holding a strigil in C1035 (Fig. 16.1d), locating metonymically the scene represented in the *palaistra*. The sharp differentiation between metaphor and metonymy is subject to theoretical criticism (Beck 1978), and both may be seen as elements of a more detailed categorization of tropes or rhetorical figures of visual expression (Groupe 1970). Nevertheless, the formal investigation of regularities in the occurrence of particular modes of connotation is a potentially important aspect of the study of figurative art.

Semantic paths

Unlike verbal communication, based on the combination of predefined expressions (ie. words), it has been noted that figurative art depends on the constant reinvention of its own expressive vocabulary (Calabrese 1985, p. 156). Each instance of a human figure produces a new expression for a given social *persona,* out of the aggregation of figure attributes. Figure types may be defined by taxonomic methods (Gordon 1981), but it should be pointed out that a single attribute may convey important meaning, even if it is not critical for the overall formal similarity between figures. In other words, it cannot be asserted that meaning is always communicated with maximal— or, indeed, predictable—redundancy. The semantic representation of compositional elements cannot, therefore, be reduced to the semantics of figure types, but should include semantic descriptions for individual attributes.

The denotative and connotative semantic units of a configuration of attributes may produce a straightforward identification of figure identity; for example, background females wearing a long-sleeved tunic and a *sakkos* headgear and holding a trinket box, are identified as slave maids (Figs 16.1b, 16.1i). Yet often the messages conveyed by a specific figure attribute may point to contradictory identifications. For example, a strigil and oil-flask identifies either an athlete or a slave boy, a *pais;* a hare appears in Classical funerary iconography either as game, identifying a hunter, or as a love present (Clairmont 1970, pp. 102–104), identifying a youth at betrothal age; the respectfulness connotation of seated stance identifies the figure as a matron, an old man, or the deceased. Often, the determination of meaning depends on the joint consideration of alternative meanings of different figure attributes.

The mechanism suggested for the resolution of ambiguity is a variant of compositional analysis using selection restrictions (Katz & Fodor 1963; Katz 1971) or settings (Eco 1976). Figure attributes, as well as types of binary compositional relations between figures, are defined as entries in a semantic dictionary, representing their possible meanings under different combinatorial and contextual selection restrictions. The fact that Classical Attic stelai provided a consistent functional setting, with well-defined pragmatics, dispenses with the need to replace the dictionary with an encyclopaedia incorporating circumstantial selection conditions (Eco 1976, pp. 98— 100). The semantic description for each dictionary entry will be a path of semantic markers, headed by its denotation but including also connotations (Fig. 16.7a). Combinatorial selection rules involve the semantic description of other attributes pertaining to the same compositional element; contextual selection rules involve the semantic description of the compositional context, ie. spatially or topologically related compositional elements. Selection restrictions distinguish between different semantic readings of the figure attribute or compositional relations.

The semantic description of compositional elements, ie. figures in stele composition, consists of the amalgamated paths of dictionary entries concerning their attributes and binary relationships, taking into account combinatorial selection rules (eg. Fig. 16.7b). Often, the co-occurrence of different attributes in a figure will allow the deletion of all readings but one, and thus semantic ambiguity will be fully resolved. However, in the general case, semantic disambiguation will require a consideration of the global iconographic context. Amalgamated paths of compositional elements and binary relationships, now involving only contextual selection restrictions, should, therefore, be included in the appropriate place of the syntactic descriptive representation, and be made available for further analysis.

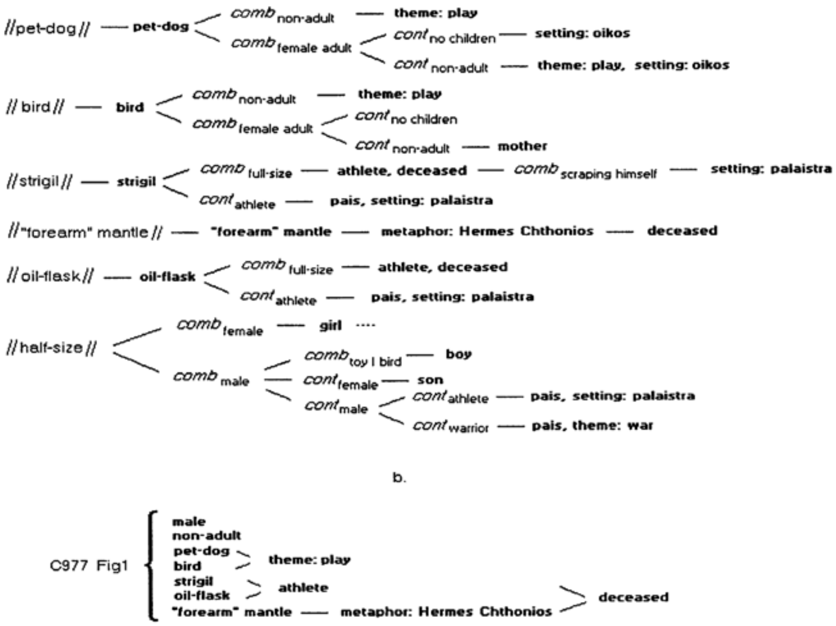


Figure 16.7 (a) Examples of semantic dictionary entries (b) Amalgamated path for the semantic representation of the standing figure in C977 (Fig. 16.1c).

Figurative grammar

An attempt is now made to integrate the analysis of stele composition in a grammatical framework. Grammars are a particular class of procedural models, involving a vocabulary of terminal symbols, a non-terminal vocabulary containing a special start symbol, and a set of production rules, ordered pairs of symbolic representations composed of members of the two vocabularies. Terminal symbols are the lexical elements empirically observed in a 'language', that is, in the class of empirically attested structures under consideration. The start symbol represents the full configuration of a structure. Other non-terminal symbols are analytic constructs, which represent syntactic categories and are used to decompose recursively the full configuration into its constituent elements. The purpose of a grammar is to generate, by applying production rules on non-terminal symbols, all (and only those) configurations of terminal symbols occurring in a particular application domain.

The syntactic component, according to a standard formulation of linguistic theory (Chomsky 1965), is distinguished into the operation of categorial (rewrite) rules and lexical insertion rules to produce a deep structure, and the subsequent operation of transformational rules on the deep structure in order to produce valid surface structures. Phonological rules are then applied on surface structures in order to produce empirically attested phonetic representations. The derivative semantic component of the grammar consists of applying a set of projection rules and a dictionary of semantic descriptions on syntactic deep structure. The capability of the grammar to produce more than one surface structure from a single deep structure explains why different phrases may share the same meaning.

Syntactic structure

The grammatical analysis of figurative art involves the definition of a categorial component, a set of fundamental syntactic categories distinct from semantic representation. In stele composition, deep structures are produced by applying rewrite rules involving syntactic categories and by replacing atomic syntactic categories with lexical compositional elements, ie. human figures. Surface structures, full symbolic descriptions of scenes on stelai, are generated by applying transformational rules to rearrange lexical elements according to their syntactic description. Since figurative art lacks a second level of articulation involving elementary visual forms devoid of meaning, there is no equivalent to the distinct phonological component of verbal language. Actual grave-reliefs are generated from surface structures by the application of an analogical, extra-semiotic process operating directly on the continuum of sculptural form.

It has been noted that, in the theory of formal languages, 'grammars of various types are equivalent to 'acceptors' (ie. automata) of various types' (Rosenfeld & Kak 1982, p. 317). The non-deterministic finite automaton presented above (Fig. 16.5) may thus be re-expressed as a grammar. To define its categorial component, consideration should be given to the need for simplicity and generality. It is, therefore, not enough to enumerate in the categorial component compositional elements used for scene expansion, such as 'foreground figure in front of base', but it is necessary to abstract their pertinent syntactic features. The grammar should also involve relational compositional structure, as shown by the descriptive and analytical formalisms defined above.

The base figure of the non-deterministic finite automaton (Fig. 16.5) is by definition a mandatory element of stele composition. This main figure (MF), characterized as the element representing the deceased, is defined as a syntactic category of the grammar. The MF is syntactically identified by means of traits such as seated stance, central or left position in the image, and receiving the visual attention of other figures; in a certain sense, it is the *topic* of the composition.

In several example scenes (Figs 16.1d, 16.1h, 16.1i), half-size figures are attached, by means of superimposition, left-to-right adjacency or contact, to the MF. They, therefore, define a further syntactic category (HF). Such HFs are clearly subordinate to the MF to which they are attached. The syntactic description of the dominant main figure may therefore be said to depend on the definition of a non-terminal main figure group (MFG) syntactic category, that always includes an MF but may also optionally contain an HF.

Expansion elements have been defined above in relation to the MF. This relation ('in front', 'behind') involves the joint consideration of left-to-right sequence and MF orientation. Plane also differentiates expansion elements into foreground and background. Since these attributes are positional, their place is not in the definition of further elements of the categorial component used to produce deep structure; they should, instead, be used to constrain transformations, regulating the rearrangement of compositional elements in order to produce valid surface structures.

Several example scenes (Figs 16.1e–16.1l) display two full-size figures in opposition, related by visual, mediated or tactile contact. Although this compositional segment is defined at the purely syntactic level, it is particularly important for the interpretation of the scenes; in fact, for interpretation purposes, stelai have been sharply bisected in the past to those which display figures in contact, usually handshake, and those which do not (Johansen 1951). The two figures in contact may or may not be adjacent, without that affecting the strength of their perceived relationship. This contact group (CG) is defined as a non-terminal syntactic category of stele composition.

When a CG is present, it always contains a MFG. The figure involved in opposition or contact with the MFG within the CG belongs to the contact figure (CF) syntactic category. On the other hand, a full scene (S)—the start symbol of the categorial component—may contain, apart from a MFG or a CG, additional figures (Figs 16.1f, 16.1h–16.1l). These figures, having clearly an accessory compositional status, are assigned to a further syntactic category, that of a secondary figure (SF).

This simplified grammar of stele composition consists, therefore, of a number of syntactic categories (Fig. 16.8a) and rewrite rules (Fig. 16.8b), producing deep structures after the replacement of terminal nodes by lexical elements. The syntactic representation of actual scenes takes the form of a phrase marker, illustrating the operation of rewrite rules on syntactic categories (Fig. 16.9). Parsing actual scenes to produce their deep syntactic representation is a process of compositional element subcategorization: figures in the scene are paired with MF, CF, SF or HF symbols, according to their compositional attributes and relationships. A set of pairings identifies a unique deep structure representation. Cases of referential ambiguity, eg. when the MF cannot be syntactically told apart from the accompanying CF, lead to the existence of more than one deep structure representation for a scene. Semantic or extra-pictorial (eg. epigraphic) evidence may then be used for resolving the ambiguity.

In general, transformations are operations on subtrees of the phrase marker. In a pure grammar, they will involve consideration of solely syntactic features of the lexical elements (ie. compositional attributes of figures recorded in the descriptive syntactic representation of the scene). An example of such a transformation concerns figure transpositions in scenes containing a CG and an SF: the SF is transposed to the foreground behind the MF (Fig. 16.8c). If the MF is seated, another transformation may take place: the SF may be transposed instead to the background within the CG, ie. between the MF and the CF (Fig. 16.8d). Another common optional transformation is image inversion.

The surface structure is an important component of a grammar of this form, since it incorporates syntactic information about 'word order' (ie. image directionality, left-to-right sequence and other compositional attributes of image elements). Although transformations are recursive, practical limits must be imposed on the number of recursions allowed; for instance, transformations producing by recursive application adjacent SFs tend in practice to be applied only once. Besides, since the production of surface structures by means of transformations is constrained by probabilistic concerns (not all transformations are equally likely), parameters may be introduced to the operation of rules.

It is important to note that, in the grammar sketched out here, the determination of surface structure from deep structure is based upon consideration of syntactic attributes, independent from semantics. While the definition of surface structure implies a theory of syntactic well-formedness, the definition of deep structure implies a theory of semantic interpretation.

Semantic interpretation

The purpose of semantic interpretation is to determine 'what is going on' in a scene, resolving ambiguity, and using available knowledge about its syntactic deep structure and the meaning of compositional elements. In order to achieve this purpose, an

(a)

Category	Description
S	Scene (start symbol)
CG	Contact group
MFG	Main figure group
MF	Main figure
HF	Half-size figure
CF	Contact figure
SF	Secondary figure

(b)

Rewrite rule		
S	→	MFG | CG
		S + (SF)
CG	→	MFG + CF
MFG	→	MF + HF
MF	→	Δ
HF	→	Δ
CF	→	Δ
SF	→	Δ

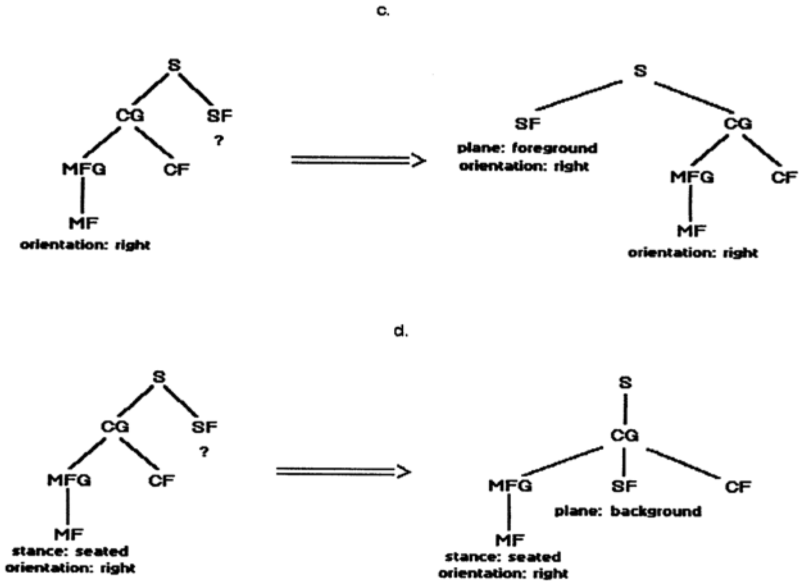


Figure 16.8 A grammar for stele composition (a) Categorial base (b) Rewrite rules (c, d) Example transformation rules regulating the transposition of secondary figures (SF) in the context of a contact group (CG).

internal representation must be defined for the entire scene, based on the integration of deep structure with semantic information.

The semantic description of figures was defined as a set of paths, alternative ‘readings’ consisting of semantic markers—concerning both figure attributes and relations with other figures—and contextual selection restrictions. To achieve integration with syntax, semantic descriptions may be incorporated as labels of terminal nodes of the syntactic phrase marker. The semantic interpretation of stele composition may thus be regarded as a cyclic upward operation through phrase markers of the grammar introduced above. Contradictory readings will be deleted, according to the operation of contextual selection procedures. Semantic markers concerning compositional relations between dominated nodes at the current level will be deleted and transferred to the dominating non-terminal node. Additional semantic markers, implicit in the contextual selection restrictions, may be transferred to node descriptions (cf. Weinreich 1971, pp. 318–319). Thus, semantic interpretation will not only define a semantic description for non-terminal syntactic categories from terminal nodes, but will also modify appropriately the semantic description of dominated nodes. The process of semantic interpretation defines, therefore, a state space, whose initial state is a phrase marker with only terminal nodes assigned a semantic description, and its final state a phrase marker with all nodes assigned a semantic description.

A simplified example illustrates why transferring markers from contextual selection restrictions is necessary for word-sense (semantic) disambiguation. On account of figure attributes alone, the full-size figure of C1035 (Fig. 16.1d) is identified as a male subadult; the smaller figure is identified as a *pais,* on account of being half-size and holding a strigil (Fig. 16.10a). In semantic interpretation, the path for the full-size figure shows that, in the context of a *pais* or on his own, the figure may be identified as the deceased. The selection restriction is not only compatible with the actual context, but is explicitly satisfied by the prior identification of the half-size figure as a *pais;* thus, the youth is marked as the deceased. Besides, the only available semantic path of the half-size figure suggests a contextual restriction of ‘athlete’; this marker is, therefore, transferred to the semantic description of the nude youth. This formalization of semantic interpretation corresponds to the intuitive

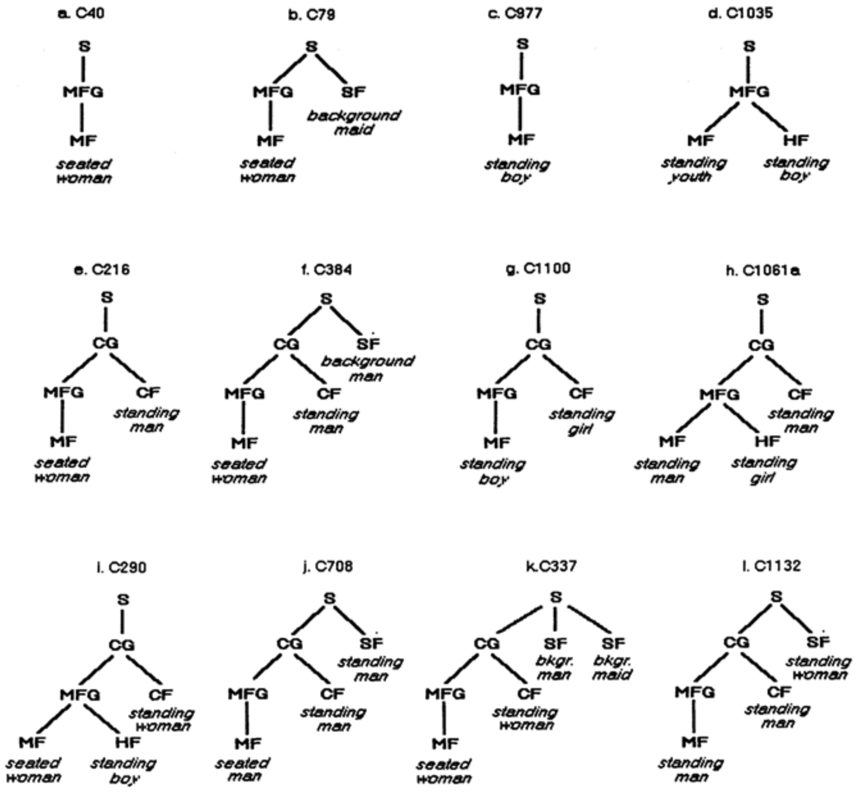


Figure 16.9 Deep structure of stele composition for examples illustrated in Fig. 16.1, according to the grammar of Fig. 16.8.

understanding of the scene. It accounts for the fact that a 'tool kit', such as the set (nudity, strigil, oil flask, bunched mantle), is often divided in stele representations between two figures dominated by the main figure group (MFG) defined above.

Besides, semantic interpretation may be used to resolve referential (syntactic) ambiguity. Especially in the presence of a contact group (CG), a scene may correspond to two different deep structures, depending on which compositional element is identified as the main figure (MF). For example, in C384 (Fig. 16.1f) either figure in handshake may be the MF, since one is seated (Fig. 16.9f), but the additional background figure looks at the other. Before semantic interpretation, the seated figure is marked as a matron, on account of seated stance and the accompanying pet dog indicating a domestic setting in the *oikos*, the standing foreground figure is marked as a young athlete, since he holds strigil and oil-flask, and the background figure merely as adult male. In deriving the semantic description of the MFG, the path of the athlete involving a contextual selection restriction of 'father or mother' is selected. Therefore, the standing athlete may be marked as deceased—thus, the MF —the seated matron as mother, and the CG, apart from the message of family unity, due to the handshake, as the theme of *mors immatura*. In amalgamating the semantic description of the MFG with the background male SF, the latter may be marked as father, and the *mors immatura* theme raised to the top level S of the phrase marker (Fig. 16.10b). Because the young man is identified as deceased, the alternative deep structure, marking the seated woman as the MF, fails to produce a complete semantic interpretation.

Semantic interpretation, as presented here, is based on normative selection restrictions. In practice, it is reasonable to accept that a semantic system functions according to probabilistic rules, so that different sources of information are weighed together in order to produce new meaning. If the frequency of actual configurations is seen as a pertinent trait of stele composition, syntax also becomes a probabilistic system. Although probability introduces complexity and involves important theoretical questions not discussed here, it is not excluded by the grammar framework presented above. *Semantic preference*, for instance, may be accounted for by allowing selection restrictions to be associated with goodness numbers; alternative interpretations could then be examined and the best selected for further processing (Charniak & McDermott 1985, p. 241).

Semantic interpretation is complicated by two additional issues. The first is related to the fact that, like verbal language, images may sometimes be deliberately ambiguous (Weinreich 1971, p. 311). For instance, in Classical Attic stelai, the identity of the deceased and other aspects of social characterization could have been determined by means of a name inscription, and thus need not have been signified in the image itself. The second issue concerns deviations from grammaticality. Since selection restrictions are used to enforce semantic well-formedness, it is possible that no global semantic description of the scene (ie. of the S node) is produced after deletion of contradictory paths. Equally, no global

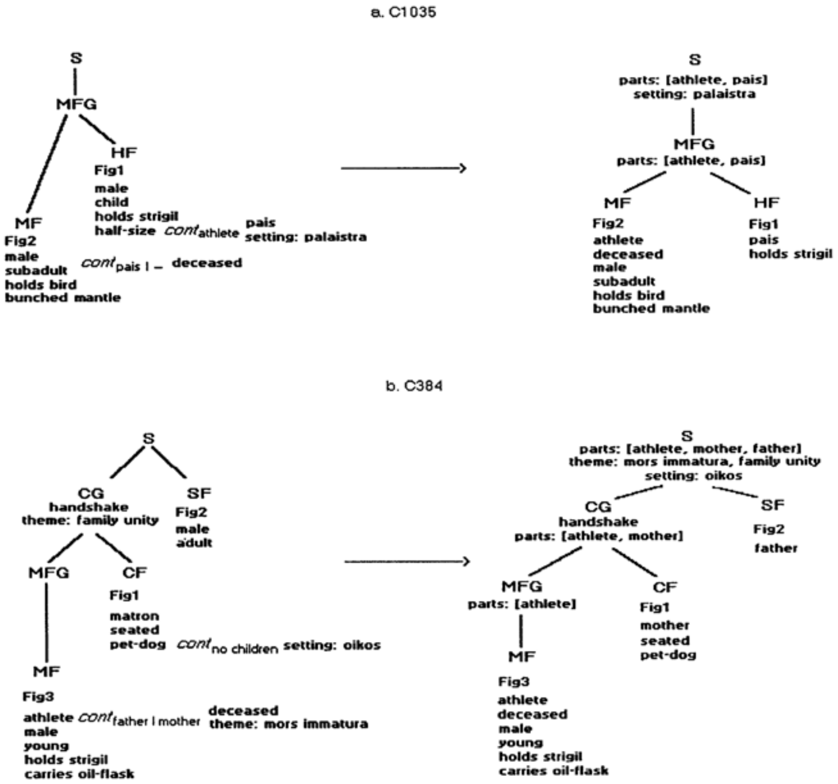


Figure 16.10 Semantic interpretation (a) Resolution of word-sense ambiguity in C1035 (Fig. 16.1d) (b) Resolution of referential ambiguity in C384 (Fig. 16.1f; cf. Fig. 16.8f).

semantic description will be produced if the image is syntactically ungrammatical, ie. is not assigned a valid deep structure representation; this is, for example, the case of exceptional compositions depicting two seated figures (eg. C781, C782).

Both issues are addressed, albeit not solved, by the proposed formalism. Intentional ambiguity will be identified by the existence of more than one final state in the state space of semantic interpretation, even after the application of contextual selection restrictions. Until adequate external evidence (eg. name inscriptions) is summoned to select among alternative interpretations, the set of final states is a correct representation of archaeological knowledge on the significance of such examples. On the other hand, using the phrase marker to represent image semantics allows the production of at least a partial interpretation of all images. Problematical images may, thus, be identified and subjected to further examination. The critical issue, in this case, is that semantic interpretation should be successful in assigning meaning to the great majority of scenes.

To determine the meaning of scenes, semantic interpretation relies heavily on the semantic dictionary. Dictionary entries presented above were of a sketchy nature; this, however, should not obscure the importance of the dictionary in a fully-fledged semantic interpretation system. The semantic dictionary is, in fact, an explicit iconology, ie. a formal theory of iconographic signification, which, unlike theories implicit in traditional discourse, is amenable to empirical confirmation.

The conjunction of the semantic dictionary with deep structures identifies semantic schemata, mappings of sets of semantic units on syntactic phrase markers. For example, *mors immatura* could be defined as an important semantic schema of stele composition, involving the depiction of a young deceased with her or his mourning parents. Other examples are *female adornment*, *sport*, and *family unity*, these schemata may be recognised as *topoi*, rhetoric figures of the content of stele composition, with known counterparts in Classical Attic funerary epigram.

Schemata are models, related to the semantic description of specific scenes in the same way as syntactic deep structures are related to surface structures; they, therefore, represent the structural—as against the contingent—component of stele semantics. Semantic schemata may be formally represented as reduced versions of phrase trees, with nodes labeled by sets of pertinent semantic markers. They may be identified in actual scenes by matching their tree description with subtrees of the phrase markers representing scenes. This method allows the identification of a schema ignoring irrelevant syntactic and semantic information. For example, *female adornment* may match all scenes with semantic markers for a female MF and a trinket box or mirror, possibly held by an SF, irrespective of further iconographic elements.

Scenes may be seen as abbreviated expressions, which, through selection of pertinent traits, stand for a presupposed narrative. This presupposition underlies the study of central problems of stele iconography, asked by Classical archaeological scholarship: for instance, the identification of the subject-matter of scenes as *leave-taking of the deceased*, *visit to the grave* or *reunion in afterlife*. Semantic schemata may also be used to represent the narrative content of stele composition. The visit to the grave subject (cf. Thimme 1964) would be defined as a semantic schema containing an isolated MF, a depicted grave-marker, and a funerary ribbon or an offering basket associated with an SF. Matching semantic schemata against the database of semantic descriptions allows theories to be explicitly examined against empirical evidence; empirical results, in their turn, allow the refinement of semantic theories of stele composition.

The symbolic representation of figurative syntax and semantics by means of phrase markers facilitates the examination of different hypotheses on the structure and content of stele composition. Example grammars, such as the one illustrated above, may be tested against empirical data, expressed according to the descriptive symbolic representation introduced above. Since they can be empirically confirmed or refuted, such grammars equate to theories of stele composition. The grammar approach presented in this study allows, therefore, an *effective,* as against an *intentional,* definition of compositional structure (Boudon 1968).

Logic programming

The Prolog programming language is particularly well-suited to problems involving complex data structures, such as lists, trees and graphs. Apart from explicit data, considered to be unconditionally true *(facts),* Prolog allows the definition and storage of implicit information about properties of the data, in the form of logic clauses *(rules)*. The Prolog query mechanism is considered to be goal-directed, since it involves not only direct pattern-matching effects, but also logical inference using all available rules in the database. The ability of Prolog to combine sophisticated data representation formalisms with goal directed programming has established it as one of the dominant computer languages for artificial intelligence research.

The structural framework for stele composition, as presented in this study, lends itself to formalization with Prolog. In the course of the project, basic data on the composition of a large number of Classical Attic stelai were extracted from a relational database, written to an ASCII file in a suitable format, and then read into the Prolog interpreter as facts. In the current implementation of the database, which will evolve as further experience is gained, each stele is represented as a Prolog structured object of the form

```
st(CatNo, TectonicType, FigList).
```

Image composition is defined as a list, FigList, which represents the string of human figures, shown from left-to-right in the image field. Apart from syntactic attributes discussed above, three major semantic markers, gender, age category and identification of the figure by inscription as deceased or survivor, are included. Members of FigList are defined as structured objects of the form

```
fig (Gender, Age, DeceasedStatus, Size, Stance, Orient, Frontal, (Gesture | GList), RelList).
```

or, optionally, to represent missing figures when a stele is partially preserved, Prolog variables of the form

```
LeftSideOfStele435.
```

The last component of the figure representation, RelList, contains binary syntactic relations of each figure with other figures; in order to facilitate queries according to both the type of relation and the identity of the other figure involved, RelList members are defined as relations with arity 2, of the form

```
rel (PatientFig, RelType).
```

The effect of this author's inexperience in Prolog programming was mitigated by the modular, forgiving nature of the language. Since the data representation chosen for stelai depends heavily on lists, an extensive number of readily available list manipulation procedures (Bratko 1986) was adopted and adapted as necessary. Selector relations were written to facilitate queries on data embedded into the complex image data structure. Using a typical selector relation, to retrieve the catalogue number and left-to-right ordinal position of standing figures, the query

```
stance (CatNo, OrdPos, standing).
```

may be issued, instead of

```
st(CatNo, _, FigList), member (fig (_,_,_,_, standing, _,_,_,_), FigList).
```

The data reduction transformations involving the merging of inverse images and the deletion of half-size and background figures, and the syntactic analysis functions concerning spatial and topological relations such as symmetry, were implemented as Prolog rules. These higher level rules access stele composition facts by invoking primitive list manipulation procedures, such as *conc*, *member* and *insert*. An example of a simple rule, identifying a ‘subimage’, an image whose composition is a segment of another composition, matching on all figure attributes specified, is:

```
subim(SubNo, No): st(SubNo,_,SubFigs), /* the composition of the two stelai st(No,_,Figs) /* is such that
conc(SubFigs,_,T), /* the figure list of the first conc(_,T,Figs), /* is a sublist of the second different
(SubFigs,Figs). /* actually a proper sublist
```

In the present state of the project, a naive implementation of a bottom-up parser for the grammar introduced above has been attempted, using a small number of examples and unary semantic unit definitions. Potentially useful rules were asserted and retracted freely, to determine the effect of alternative operations on image composition.

Formalization with Prolog is still at the experimentation stage, yet a few important points already emerge. First, the Prolog instantiation mechanism provides a simple and powerful solution for handling missing data. Such data are expressed as Prolog variables, so that data structures can be compared either for optimistic matching (ie. assuming that the missing attributes would also match) or, when deemed necessary, for literal equality (Bratko 1986, pp. 168–9). The practical advantages derived from such flexibility cannot be underestimated. Missing data are very common with Classical Attic stelai, as with most archaeological evidence.

Partially preserved stelai are a case in point: named dummy variables were inserted into the data representation as the first and/or last members of the figure list, to indicate that further figure (s) may have been originally included on the broken left or right side. Thus, when searching for a certain compositional type, it is possible to find not only fully preserved stelai that match, but also partially preserved stelai. In iconographic research, querying the database has an heuristic role, and it is important that as much relevant information as possible is extracted from the data.

Second, the Prolog database, modelling implicit information about the structure of figurative ‘texts’ in the form of rules, is a potentially very useful tool for data exploration. Clauses may be combined in queries to extract information according to complex criteria: for instance, all expansions of an image by means of half-size figure addition and image inversion; all mirror images of segments of a given image, perhaps restricted according to attributes of figures depicted; or, full images or segments of images that display an internal symmetric structure so far as figure stance, plane, orientation and degree of frontality, or an arbitrary combination of figure attributes is concerned. In fact, all analytical syntactic functions, implemented as Prolog clauses, could be used in conjunction to produce the desired result.

Third, the declarative nature of Prolog makes it suitable for the expression and testing of grammars. Grammars can be used as production systems, to examine the set of possible structures they generate; alternatively, symbolic representations of actual images, expressed as Prolog facts, can be used as input to a grammar to determine if it accounts adequately for the empirical evidence. In the latter sense, the system will function as an expert system, generating inferences about the syntactic structure and semantic interpretation of an image, on the basis of its symbolic description. With a full implementation of Prolog, offering definite clause grammar support, syntactic rewrite rules and transformations could be written directly as Prolog rules, a fact simplifying greatly the process of parsing.

In attempting to implement more complex aspects of the model of figurative art introduced above, the computing needs of the project increase. On the one hand, adequate memory space is necessary, in order to allow the storage of Prolog facts representing a large amount of empirical data; computational efficiency also becomes important as the complexity of the symbolic representation and the number of rules increases. On the other hand, ensuring both the correctness and performance of the system as its complexity increases becomes a problem. Apart from powerful hardware, mechanisms for forward-chaining and conflict resolution will be necessary in order to tackle adequately aspects of grammatical inference and semantic interpretation. True defaults are needed to deal properly with missing information. Inheritance and facets should be useful for constructing the semantic dictionary; neural networks may be an interesting alternative for the representation of global semantic information. Support for conditional probability and abduction would allow a more realistic simulation of the way meaning is assigned to images. Some at least of these facilities, normally available as part of expert systems, will be necessary in the next stage of the project. It is hoped that they will be found in the form of a programmable toolkit, allowing the flexibility in data representation currently enjoyed with Prolog.

Conclusion

A central problem in the formal study of figurative art in archaeology concerns its symbolic representation. In devising such a symbolic representation, it is necessary to acknowledge explicitly the semiotic nature of images, and the importance of structure and context on meaning. On the other hand, far from providing license for the uncritical adoption of linguistic terminology, a structural analysis standpoint makes it all the more necessary to devise a rigorous formal calculus that deals with the special nature of figurative 'texts'.

The composition of Classical Attic stelai, a homogeneous, chronologically bounded series of funerary markers from a limited geographic area, was used as a case study. Alternative approaches to image description and analysis were tried, ranging from a componential descriptive model based on global image attributes, to a grammar with separate syntactic and semantic components. Logic programming was used for the practical implementation of operational models. Although it was clear that no formal system for the analysis of figurative art can be successful if it does not take into account the morphological and functional specificity of the particular archaeological problem at hand, a few general points did emerge.

Since both image parts, and spatial and topological relationships between parts, are important aspects of image structure, labeled relational graphs (including strings and trees) are more appropriate for the symbolic representation of images than the rectangular data matrix into which data are usually shoehorned for quantitative analysis. Besides, procedural models, such as those based on transformation operations linking image composition states, are more succinct and powerful than their descriptive counterparts (ie. classifications). Analytical syntactic functions—in effect procedural extensions to the proposed symbolic representation of stele composition—provide a powerful tool for the extraction of latent compositional information. Finally, grammars, a special class of procedural models, seem to be particularly promising for the formalization of figurative art.

In the grammar presented above, syntax and semantics were seen as distinct, but complementary components of stele composition, the former providing a context of discovery, the latter a context of validation. Even the sketchy construction of a pictorial grammar involves a number of non-trivial considerations: defining a categorial component appropriate to the substantive domain; developing a unified formal representation for deep structure in the form of a phrase marker; defining the form of a global semantic dictionary; amalgamating semantic description of compositional elements in a way appropriate to visual sign-functions; dealing with different fields and tropes of connotative meaning; providing for effective retrieval of latent pattern, both syntactic and semantic; and, not least, demonstrating how syntactic deep structure can operate on ambiguous image parts to produce an unambiguous global semantic representation.

The mechanisms of grammaticality and disambiguation, responsible for morphogenesis and the generation of meaning respectively, are fundamental for the understanding of images found in the archaeological record. In traditional archaeological discourse, the discussion of figurative art is as a rule highly sophisticated, involving a complex web of arguments on the internal formal structure, the functional context and the significance of images. Formal and quantitative archaeology, on the other hand, has a good track record in the field of morphology and classification. If formal archaeology is to deal adequately with the analysis of figurative art, it must extend its theoretical arsenal beyond classification, to encompass a more powerful data description and analysis framework. If the study of iconography is to achieve the generation of verifiable theories, it requires formalization. For these reasons, the combination of structural analysis and semiotics with formal methodology has been seen to be a fruitful approach for the study of figurative art.

Acknowledgements

A large part of this chapter was based on work conducted in 1987, during an Honorary Junior Research Fellowship in the Research Centre for Computer Archaeology at Stafford Polytechnic and reflects to a significant extent my concerns at that time. I wish to thank Kathy Baker for initial exchanges on Prolog, expert systems and knowledge representation. The present chapter benefitted from comments and discussion with participants of the *Communication in Archaeology* symposia in Venezuela, especially Mike Fletcher, Marie-Salomé Lagrange and Arthur Stutt. I could not have participated in WAC2 without financial aid arranged by the organisers and editors of this volume, a fact which I gratefully acknowledge.

References

Arnold, D. 1969. *Die Polykletnachfolge.* Berlin: Walter de Gruyter.

Ashmole, B. 1972. *Architect and Sculptor in Classical Greece.* New York: New York University Press.

Barthes, R. 1966. Introduction à l'analyse structurale des récits. *Communications* **8**.

Barthes, R. 1967. *Système de la Mode.* Paris: Seuil.

Beck, B.E.F. 1978. The metaphor as a mediator between semantic and analogic modes of thought. *Current Anthropology* **19(1),** 83–97.

Birdwhistell, R.L. 1952. *Introduction to Kinesics.* Louisville, Kentucky: University of Louisville Press.

Boast, R. 1985. Documentation of prehistoric style: a practical approach. *Archaeological Review from Cambridge* **4(2),** 159–170.

Bogatyrev, P. 1938. Funkcie Kroja na Moravskom Slovensku, Turcansky Sc. Martin: Le funzioni dell'abbigliamento popolare nella Slovacchia Morava. In *Semiotica della Cultura Popolare,* Solimini, M. (ed.), 85–148. Verona: Bertrani.

Boudon, R. 1968. *A Quoi Sert la Notion de 'Structure'?* Paris: Gallimard.

Bratko, I. 1986. *Prolog Programming for Artificial Intelligence.* London: Addison-Wesley.

Calabrese, O. 1985. *Il Linguaggio dell'Arte.* Milan: Bompiani.

Carpenter, R. 1960. *Greek Sculpture: a critical review.* Chicago: University of Chicago Press.

Carpettini, G.P. 1980. *Aspetti della Semiotica: principi e storia.* Turin: Einaudi.

Chang, K.C. 1967. *Rethinking Archaeology.* New York: Random House.

Charniak, E. & D.McDermott 1985. *Introduction to Artificial Intelligence.* London: Addison-Wesley.

Chomsky, N. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: Massachusetts Institute of Technology Press.

Clairmont, C. 1970. *Gravestone and Epigram: Greek memorials from the archaic and classical period.* Mainz: Philip von Zabern.

Clark, G.A. 1982. Quantifying archaeological research. In *Advances in Archaeological Method and Theory,* Schiffer, M.B. (ed.), **5,** 217–273. New York: Academic Press.

Clarke, D.L. 1968. *Analytical Archaeology.* London: Methuen.

Conkey, M.W. 1982. Boundedness in art and society. In *Symbolic and Structural Archaeology,* Hodder, I. (ed.), 115–128. Cambridge: Cambridge University Press.

Conkey, M.W. 1989. Structural analysis of Paleolithic art. In *Archaeological Thought in America,* Lamberg-Karlovsky, C.C. (ed.), 135–154. Cambridge: Cambridge University Press.

Constantopoulos, P., Y.Drakopoulos & Y.Yeorgaroudakis 1990. Multimedia document retrieval by pictorial content. In *Multimedia Office Filing: the MULTOS approach,* Thanos, C. (ed.), 331–349. Amsterdam: Elsevier North-Holland.

Conze, A. (ed.) 1893. *Die Attischen Grabreliefs.* Berlin: Walter de Gruyter.

Dallas, C.J. 1987. *The Significance of Costume on Classical Attic Grave 'Stelai': a statistical analysis.* PhD thesis, University of Oxford, Oxford: University of Oxford.

Davies, G. 1985. The significance of the handshake motif in classical funerary art. *American Journal of Archaeology* **89,** 627–640.

Deetz, J. 1967. *Invitation to Archaeology.* New York: Natural History Press.

Diepolder, H. 1931. *Die Attischen Grabreliefs des 5. und 4.Jahrhunderts v. Chr.* Berlin: H.Keller.

Dohrn, T. 1957. *Attische Plastik vom Tode des Phidias bis zum Wirken der Großen Meister des IV. Jahrhunderts v. Chr.* Krefeld: Scherpe-Verlag.

Doran, J. 1986. Formal methods and archaeological theory: a perspective. *World Archaeology* **18(1),** 21–37.

Eco, U. 1968. *La Struttura Assente.* Milan: Bompiani.

Eco, U. 1976. *A Theory of Semiotics.* Bloomington: Indiana University Press.

Fillmore, C.J. 1968. The case for case. In *Universals in Linguistic Theory,* Bach, E. & Harms, R.T. (eds), 1–88. New York: Holt, Rinehart and Winston.

Gardin,J.-C. 1978. *Code pour l'Analyse des Ornements.* Paris: Centre Nationale de Recherches Scientifiques.

Gardin, J.-C. 1980. *Archaeological Constructs.* Cambridge: Cambridge University Press.

Glassie, H. 1975. *Folk Housing in Middle Virginia: a structural analysis of historical artifacts.* Knoxville: University of Tennessee Press.

Gordon, A.D. 1981. *Classification methods for the exploratory analysis of multivariate data.* London: Chapman and Hall.

Groupe, M. 1970. *Rhétorique Générale.* Paris: Larousse.

Hall, E. 1959. *The Silent Language.* Greenwich, Conn.: Fawcett Publications.

Hardin, M.A. 1983. The structure of Tarascan pottery painting. In *Structure and Cognition in Art,* Washburn, D.K. (ed.), 8–24. Cambridge: Cambridge University Press.

Himmelmann, N. 1971. *Archeologisches zum Problem der griechischen Sklaverei.* Wiesbaden: Steiner.

Himmelmann-Wildschütz, N. 1956. *Studien zum Ilissos-Relief.* Munich: Prestel Verlag.

Hjelmslev, L. 1943. *Prolegomena to a Theory of Language.* Madison: University of Wisconsin Press.

Hodder, I. 1982a. Sequences of structural change in the Dutch neolithic. In *Symbolic and Structural Archaeology,* Hodder, I. (ed.), 162–177. Cambridge: Cambridge University Press.

Hodder, I. 1982b. Theoretical archaeology: a reactionary view. In *Symbolic and Structural Archaeology,* Hodder, I. (ed.), 1–16. Cambridge: Cambridge University Press.

Hodder, I. 1986. *Reading the Past: Current Approaches to Interpretation in Archaeology.* Cambridge: Cambridge University Press.

Hodder, I. (ed.) 1987. *The Archaeology of Contextual Meanings.* Cambridge: Cambridge University Press.

Hoffmann, H. 1977. *Sexual and Asexual Pursuit: a structuralist approach to Greek vase painting.* London: Royal Anthropological Institute.

Hymes, D. 1970. Linguistic models in archeology. In *Archéologie et Calculateurs,* Gardin, J.-C. (ed.), 91–120. Paris: Centre Nationale de Recherches Scientifiques.

Jacobson, R. 1956. Two aspects of language and two types of aphasic disturbance. In *Fundamentals of Language,* Jacobson, R. & Halle, M. (eds), 52–82. The Hague: Mouton.

Johansen, K.F. 1951. *The Attic Grave-Reliefs of the Classical Period: an essay in interpretation.* Copenhagen: Munksgaard.

Karouzos, C. 1966. Telauges mnema. *Charisterion eis A. Orlandon* **3,** 253–280.

Karouzou, S. 1969. Der Hermes von Andros und seine Gefahrtin. *Athenische Mitteilungen* **84,** 143–157.

Katz, J.J. 1971. Semantic theory. In *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology,* Steinberg, D.D. & Jakobovits, L. (eds), 297–307. Cambridge: Cambridge University Press.

Katz, J.J. & J.A.Fodor 1963. The structure of a semantic theory. *Language* **39,** 170–210.

Lagrange, M.-S. & M.Renaud 1985. Intelligent knowledge based systems in archaeology: a computerized simulation of reasoning by means of an expert system. *Computers and the Humanities* **19,** 37–52.

Lagrange, M.-S. & M.Renaud 1988. Cas No 6: Superikon, essai de cumul de six expertises en iconographie. In *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology,* Gardin, J.-C., Guillaume, O., Herman, P.-Q., Hesnard, A., Lagrange, M.-S., Renaud, M., & Zadora-Rio, E. (eds). Chichester: Ellis Horwood.

Lakoff, G. 1971. On generative semantics. In *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology,* Steinberg, D.D. & Jakobovits, L.A. (eds), 232–296. Cambridge: Cambridge University Press.

Le Rider, G. 1975. *Code pour l'Analyse des Monnaies.* Paris: Centre Nationale de Recherches Scientifiques.

Leroi-Gourhan, A. 1965. *Treasures of Palaeolithic Art.* New York: Abrams.

Leroi-Gourhan, A. 1982. *The Dawn of European An: an introduction to palaeolithic cave painting.* Cambridge: Cambridge University Press.

Levesque, H.J. 1986. Knowledge representation and reasoning. *Annual Review of Computer Science* **1,** 225–287.

Lévi-Strauss, C. 1949. *Les Structures Élémentaires de la Parenté.* Paris: Presses Universitaire de France.

Lippold, G. 1923. *Kopien und Umbildungen griechischen Statuen.* Munich: Beck.

Lisboa, I.M.G. 1985. Meaning and messages: mapping style in the Iberian chalcolithic. *Archaeological Review from Cambridge* **4(2),** 181–196.

Llamazares, A.M. 1989. A semiotic approach in rock-art analysis. In *The Meanings of Things: material culture and symbolic expression,* Hodder, I. (ed.), 242–248. London: Unwin Hyman.

Lloyd, G. 1962. Right and left in Greek philosophy. *Journal of Hellenic Studies* **82,** 55–66.

Maclay, H. 1971. Linguistics: overview. In *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology,* Steinberg, D.D. & Jakobovits, L.A. (eds), 157–182. Cambridge: Cambridge University Press.

Miller, D. 1982. Artefacts as products of human categorisation processes. In *Symbolic and Structural archaeology,* Hodder, I. (ed.), 17–25. Cambridge: Cambridge University Press.

Munn, N.D. 1966. Visual categories: an approach to the study of representational systems. *American Anthropologist* **68,** 936–950.

Newman, W.M. & R.F.Sproull 1981. *Principles of Interactive Computer Graphics.* New York: McGraw-Hill.

Renfrew, C. 1982. Explanation revisited. In *Theory and Explanation in Archaeology: the Southampton conference,* Renfrew, C., Rowlands, M.L. & Seagraves, B.A. (eds), 5–23. New York: Academic Press.

Roach, M.E. & J.B.Eicher 1965. *Dress, Adornment and the Social Order.* Chichester: Wiley.

Rosenfeld, A. & A.C.Kak 1982. *Digital picture processing vol. 2.* New York: Academic Press.

Salomé, M.-R. 1980. *Code pour l'Analyse des Représentations Figurees sur les Vases Grecs.* Paris: Centre Nationale de Recherches Scientifiques.

Saxe, A.A. 1970. *Social Dimensions of Mortuary Practices.* PhD thesis, University of Michigan, Michigan: University of Michigan.

Schefold, K. 1952. Zur Deutung der klassischen Grabreliefs. *Museum Helveticum,* **9,** 107–112.

Schmaltz, B. 1970. *Untersuchungen zu den attischen Marmorlekythen.* Berlin: Mann.

Schmaltz, B. 1983. *Griechische Grabreliefs.* Darmstadt: Wissenschaftliche Buchgesellschaft. Small, J.P. 1987. Left, right, and center: direction in Etruscan art. *Opuscula Romana* **16,** 125–135.

Strocka, V.M. 1979. Variante, Wiederholung und Serie in Griechischer Bildhauerei. *Jahrbuch des Deutschen archeologischen Instituts* **94,** 143–173.

Stupperich, R. 1977. *Staatsbegrübnis und Privatgrabmal im Klassischen Athens.* PhD thesis, University of Münster, Münster: University of Münster.

Taylor, T. 1987. Flying stags: icons and power in Thracian art. In *The Archaeology of Contextual Meanings,* Hodder, I. (ed.), 117–132. Cambridge: Cambridge University Press.

Thimme, J. 1964. Die Stele der Hegeso als Zeugnis des Attischen Grabkults. *Antike Kunst* 7**,** 17–29.

Washburn, D.K. 1983a. Symmetry analysis of ceramic design: two tests of the method on neolithic material from Greece and the Aegean. In *Structure and Cognition in Art,* Washburn, D.K. (ed.), 38–65. Cambridge: Cambridge University Press.

Washburn, D.K. 1983b. Toward a theory of structural style in art. In *Structure and Cognition in Art,* Washburn, D.K. (ed.), 1–7. Cambridge: Cambridge University Press.

Weinreich, U. 1971. Explorations in semantic theory. In *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology,* Steinberg, D.D. & Jakobovits, L.A. (eds), 308–328. Cambridge: Cambridge University Press.

Whallon, R. 1972. A new approach to pottery typology. *American Antiquity* **37,** 13–33.

Wobst, M. 1977. Stylistic behaviour and information exchange. In *For the Director: research essays in honor of James B.Griffin,* Cleland, C.E. (ed.), 317–342. Ann Arbor: Museum of Anthropology, University of Michigan.

17

VANDAL: an expert system dealing with the provenance of archaeological ceramics, based on chemical, mineralogical and data analysis information

MARIE-SALOMÉ LAGRANGE & VANDA VITALI

Why an expert system and for whom?

VANDAL was designed for researchers or students of archaeology, requiring no prior knowledge of materials science, statistics or computer science; it is, however, advisable that a user obtain some basic information regarding expert systems prior to the utilization of VANDAL (Laurière 1986a).

It was conceived to demonstrate the potential and the limitations of technical information in archaeometric provenance studies, a task often too complex for archaeologists alone. It was subsequently developed to help archaeologists in the interpretation of results from technical (archaeometric) analysis (chemical, mineralogical, and data analysis), within the framework of archaeometric provenance studies.

On the one hand, VANDAL produces one ‘conclusion’ (or a set of ‘conclusions’) regarding ceramic artefact provenances from the data given by the user; on the other hand, it can serve as a tool for the verification/rationalization of already established conclusions, by the virtue of its nature; production rules contained in the knowledge base of the system, readable even by a non-computer specialist, can explicitly reproduce the reasoning employed by the system in the decision-making process.

Archaeometric provenance studies

To determine the provenance of an archaeological object involves relating an object to a geographic location. This location can be the place where the object was excavated or otherwise found, where it was made, or where the raw materials used in its production originated. While the excavation location of an object is generally established on the basis of archaeological recording, the determination of the place of manufacture, or the determination of the geological/geographic origin of the raw materials used in the production of the object, are questions with which archaeometric provenance studies are concerned. The excavation context alone sometimes allows us to establish that the object of interest did not ‘travel’, that is to say that the find location is the location of the object’s manufacture. However, in many cases, the place of the manufacture and/or the geologic/geographic origin of the raw materials used in the production of the object cannot be determined without technical analyses.

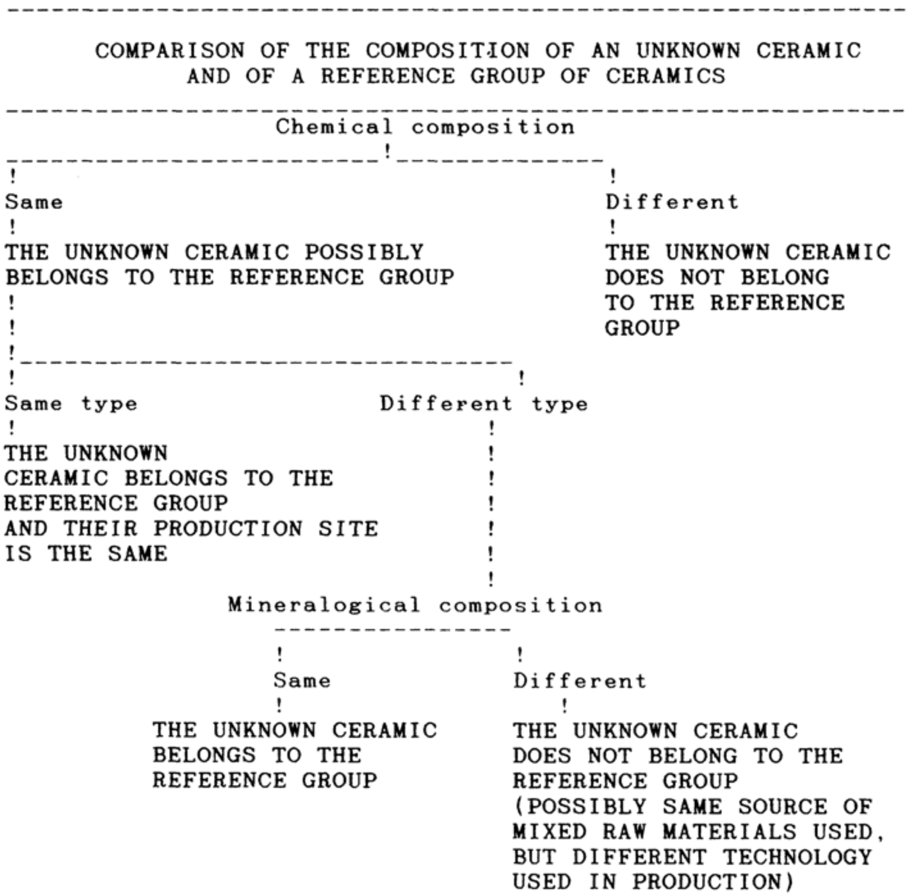
In such studies, the link between (a) an object and a centre for its manufacture, or (b) an object and the raw materials used in its production, are established on the basis of a comparison with material from which the object was made, and a reference material representing either a centre of manufacture or a raw material. The study of the chemical and mineralogical composition have been found to be the most useful techniques for characterization in current ceramic provenance studies.

Determining the provenance of an archaeological object based on archaeometric information consists, therefore, of a series of comparisons (chemical and mineralogical) between the material of an object and the reference material(s), representing either a known centre of manufacture or a known raw material (Vitali 1989).

Analysis of reasoning

The analysis of archaeological reasoning and its simulation using expert systems have been the subject of a number of studies (see in particular Lagrange & Renaud 1983; Lagrange & Renaud 1985; Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988). These studies have shown that expert systems are tools particularly suited for the rationalization and demonstration of reasoning in social and human sciences, fields in which there still exists ‘la logique du plausible’, but which demand an ever-increasing scientific rigour. VANDAL is a direct descendant of these studies having, at the same time, a strongly pronounced utilitarian character.

Within the framework of a larger study dealing with the role of technical (archaeometric) information in the reconstruction of the past (Vitali & Lagrange 1988; Vitali 1989), a detailed and critical study of reasoning in archaeometric provenance determination has been carried out.



```
------------------------------------------------------------

      COMPARISON OF THE COMPOSITION OF AN UNKNOWN CERAMIC
            AND OF A REFERENCE GROUP OF CERAMICS

------------------------------------------------------------
                  Chemical composition
                         !
-------------------------!--------------
!                                        !
Same                                     Different
!                                        !
THE UNKNOWN CERAMIC POSSIBLY             THE UNKNOWN CERAMIC
BELONGS TO THE REFERENCE GROUP           DOES NOT BELONG
!                                        TO THE REFERENCE
!                                        GROUP
!----------------------------------
!                                  !
Same type                 Different type
!                                 !
THE UNKNOWN                       !
CERAMIC BELONGS TO THE            !
REFERENCE GROUP                   !
AND THEIR PRODUCTION SITE         !
IS THE SAME                       !
                                  !
               Mineralogical composition
                  ----------------
                  !               !
                  Same            Different
                  !                  !
          THE UNKNOWN CERAMIC     THE UNKNOWN CERAMIC
          BELONGS TO THE          DOES NOT BELONG TO THE
          REFERENCE GROUP         REFERENCE GROUP
                                  (POSSIBLY SAME SOURCE OF
                                  MIXED RAW MATERIALS USED,
                                  BUT DIFFERENT TECHNOLOGY
                                  USED IN PRODUCTION)
```

Figure 17.1 Comparison of the composition of an unknown ceramic and of a reference group of ceramics.

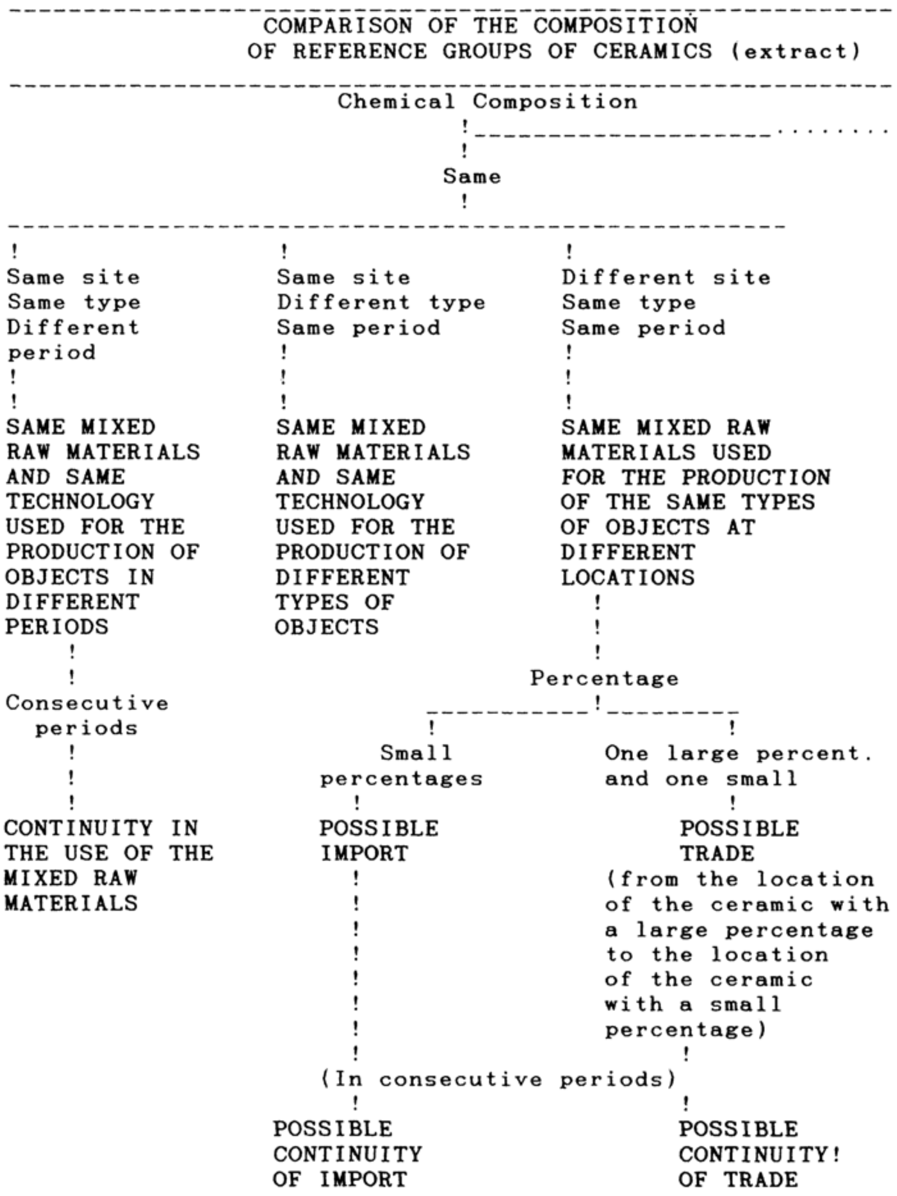
Figures 17.1 and 17.2 give examples of the schematics of the reasoning for such studies that explicitly show the comparisons that take place, and the conclusions that can be reached in archaeometric provenance determinations. It should be noted that although the comparisons are based on chemical and mineralogical criteria, the questions are asked in archaeological terms: the reference ceramics are described by their excavation or find site (SITE), by their type (TYPE), by the chronological period to which they belong (PERIOD) and, sometimes by the percentage of finds.

Archaeometric information forms the basis for all comparisons. This information is presented as the *same* or *different* composition, a differentiation made by experts. Therefore, for the reader of Figures 17.1 and 17.2, archaeometric information must be known and considered correct. Figures 17.1 and 17.2 explicitly state the possible conclusions that can be drawn from archaeometric information. These conclusions are of archaeological significance. For example, if two ceramic reference groups, from the same site, of the same period but of different type, have different chemical compositions but the same mineralogical composition (see Figure 17.2), then it can be said that they were made using the same mixture of raw materials, but by a different technology of production. Similarly, if an unknown ceramic and a reference group of ceramics have the same chemical composition, and are of the same type, then it can be deduced that the unknown ceramic belongs to the reference group, and that their place of manufacture was the same. Also included are possible, but less likely, conclusions that are valid only in certain cases, such as the deductions made on the basis of percentage of finds. These conclusions, of course, are clearly indicated as possibilities only.

Strategy for the implementation of VANDAL

The strategy adopted for reproducing the reasoning in Figures 17.1 and 17.2 in the form of an expert system, was to write production rules which compare the chemical and mineralogical composition of pairs of objects (unknown ceramics, reference ceramics or raw materials) in the following manner:

- unknown ceramic (or group of ceramics) to reference group of ceramics;
- reference group of ceramics to reference group of ceramics;
- unknown ceramic or reference group of ceramics to raw material constituents.



```
----------------------------------------------------------
                 COMPARISON OF THE COMPOSITION
                OF REFERENCE GROUPS OF CERAMICS (extract)
----------------------------------------------------------
                      Chemical Composition
                              !___________________........
                              !
                             Same
                              !
---------------------------------------------------
!                  !                  !
Same site          Same site          Different site
Same type          Different type     Same type
Different          Same period        Same period
period             !                  !
!                  !                  !
!                  !                  !
SAME MIXED         SAME MIXED         SAME MIXED RAW
RAW MATERIALS      RAW MATERIALS      MATERIALS USED
AND SAME           AND SAME           FOR THE PRODUCTION
TECHNOLOGY         TECHNOLOGY         OF THE SAME TYPES
USED FOR THE       USED FOR THE       OF OBJECTS AT
PRODUCTION OF      PRODUCTION OF      DIFFERENT
OBJECTS IN         DIFFERENT          LOCATIONS
DIFFERENT          TYPES OF             !
PERIODS            OBJECTS              !
    !                                   !
    !                                Percentage
Consecutive                   ___________!_________
  periods                     !                    !
    !                       Small          One large percent.
    !                   percentages        and one small
    !                       !                      !
CONTINUITY IN           POSSIBLE                POSSIBLE
THE USE OF THE          IMPORT                  TRADE
MIXED RAW                 !                (from the location
MATERIALS                 !                of the ceramic with
                          !                a large percentage
                          !                to the location
                          !                of the ceramic
                          !                with a small
                          !                percentage)
                          !                      !
                        (In consecutive periods)
                          !                      !
                       POSSIBLE                POSSIBLE
                       CONTINUITY              CONTINUITY!
                       OF IMPORT               OF TRADE
```

Figure 17.2 Comparison of the composition of reference groups of ceramics.

These comparisons produce one or more conclusions related to the provenance of ceramics, according to criteria such as those listed in Figures 17.1 and 17.2, in which the user supplies a set of ceramic descriptions. VANDAL examines the list and establishes the necessary comparisons. The user then introduces the values for the comparison, based on the findings of technical analysis (eg. object 1 has the *same* chemical composition as object 2, or object 1 has a *different* chemical composition from object 2). The expert system examines the added information and produces conclusions.

Thus, the procedure to be followed is to:

1. supply the information to the expert system (description of the ceramics and raw materials in the form of an initial 'fact base', as well as the comparison values once the comparison pairings have been established) in a format specified by the SNARK shell.
2. apply to the data the rules that constitute the knowledge of the VANDAL 'expert' (rules of chemical and mineralogical comparisons and rules that produce archaeological conclusions). All relevant rule bases are called in succession, and the system is requested to execute them one after the other, starting from an initial fact base, and each time accumulating results.

1	G2-UNK-CERAM-1	ROLE	UNK-OBJECT
2	G2-UNK-CERAM-1	TYPE	GEOMETRIC
3	G2-UNK-CERAM-1	SAME-CHEM-COMPOS	G2-REF-CERAM-2
4	G2-REF-CERAM-2	ROLE	REF-OBJECT
5	G2-REF-CERAM-2	TYPE	GEOMETRIC

Figure 17.3 Example of a fact base.

RULE:	BELONG-UNK		
SI	(1)	ROLE	(A)=UNK-OBJECT
	(2)	TYPE	(B)=REF-OBJECT
	(3)	TYPE	(A)=TYPE(B)
	(4)	SAME-CHEM-COMPOS	(A)=(B)
ALORS			
		BELONG	(A) <== (B)
FR			

Figure 17.4 Example of a production rule.

A brief overview of SNARK

VANDAL is implemented in SNARK, an expert system shell developed by Laurière (Laurière 1986b; Lagrange & Renaud 1988; Vitali & Lagrange 1988). It is written in Pascal, uses forward chaining reasoning, is non-monotonous, and has a capacity to function with a logic of order 0, 1, and/or 2.

The principal reason for choosing SNARK, apart from its well known robustness and its flexibility, was its capacity to handle variables (first order logic) and second order logic. It is not necessary for a user of VANDAL to know in detail the various modes of functioning, but it is important to recognize that these qualities are rarely found in expert system shells currently available, and that they were heavily used in designing VANDAL. SNARK is now commercially available in various versions, on mainframes and personal computers. To use VANDAL, beside the VANDAL files, it is necessary to have access to an IBM compatible personal computer (640K memory).

SNARK functions by applying production rules (interpretive rules) to the facts provided by the user (see Figure 17.3 and appendix). These rules cannot be applied unless their antecedents are unified with the facts of the fact base. Each time a new fact is produced by the launching of a rule whose antecedents are satisfied, it is added to the initial or the current fact base. This new fact, in turn, allows an application of another rule, and so on.

The inference engine unifies antecedents of rules and facts by attempting to instantiate (substitute or replace) variables by the entities or values given in the fact base (see Fig. 17.3).

Figure 17.3 displays the following elements:

1. G2-UNK-CERAM-1 is an unknown object, ie. an unknown ceramic.
2. Its type is geometric.
3. It has the same chemical composition as G2-REF-CERAM-2.
4. G2-REF-CERAM-2 is a reference group of ceramics.
5. Its type is geometric.

Figure 17.4 shows a production rule which the system can apply to the fact base shown in Figure 17.3.

Rule BELONG-UNK in Figure 17.4 means:

If (A) is an unknown ceramic, if (B) is a reference group of ceramics, if (A) and (B) are of the same type, and if (A) has the same chemical composition as (B), then (A) belongs to the reference group (B).

In this example, antecedent (1) of the rule base can be matched with the line (1) of the fact base because their relationship (ROLE) and the value (UNK-OBJECT) are the same. Consequently, variable (A) is a candidate for an instantiation by the entity G2-UNK-CE RAM-1. In the case of antecedent (2), the same applies in the case of fact (4). Thus variable (B) becomes a candidate for an instantiation by G2-REF-CERAM-2. Antecedents (3) and (4), if facts (5) and (3) are considered, are compatible with envisaged instantiations. The rule can be applied and the action of the rule produces a new fact according to the statement:

G2-UNK-CERAM-1 BELONG G2-REF-CERAM-2

This ‘final fact’ (conclusion) will then be added to the current fact base.

Conclusions

Expert systems are widely used nowadays in industrial or technical contexts to capture, and render repeatable and usable, all kinds of implicit knowledge, which cannot be automated through algorithmic computer programming. The ability of expert systems to represent qualitative or even fuzzy knowledge, and their ease of use so far as computer non-specialists are concerned, are assets largely compensating for their relatively meagre success, given the processing time and memory requirements.

It is probable that the difficulties we encountered in the proper integration of technical disciplines into archaeology may very well reflect a general problem. The construction of VANDAL has proved of considerable use in analysing, and explicitly rendering, all possible interpretive statements using archaeometric information to produce archaeological conclusions, in the case of ceramic provenance. We have designed and implemented a research and tutoring aid to archaeologists, in the shape of a users’ manual which includes the necessary computer files on disk (to be published).

Appendix

Example of a VANDAL *initial fact base*

(For ceramic reference groups)

G2-REF-CERAM-3	NAT	CERAMICS
G2-REF-CERAM-3	ROLE	REF-OBJECT
G2-REF-CERAM-3	PERCENT	7
G2-REF-CERAM-3	CONSEC	MC2
G2-REF-CERAM-3	TYPE	CORDED
G2-REF-CERAM-3	SITE	GEOKSYUR
G2-REF-CERAM-3	PERIOD	MC1 (middle chalcolithic)
G2-REF-CERAM-3	NUM	3
(For raw materials)		
G2-RAW-MAT-32	ROLE	RAW-MATERIAL
G2-RAW-MAT-32	SITE	YAZ-DEPE
G2-RAW-MAT-32	NUM	32
G2-RAW-MAT-32	NAT	CLAY
G2-RAW-MAT-33	ROLE	RAW-MATERIAL
G2-RAW-MAT-33	SITE	NUR-TEPA
G2-RAW-MAT-33	NUM	33
G2-RAW-MAT-33	NAT	TEMPER
(For unknown ceramics)		
G2-UNK-CERAM-34	ROLE	UNK-OBJECT
G2-UNK-CERAM-34	ASSOCIATION	ST-SIMON
G2-UNK-CERAM-34	NUM	34
G2-UNK-CERAM-34	NAT	CERAMICS
G2-UNK-CERAM-34	TYPE	YELLOW

Extracts from a* VANDAL *rule base

★ EXAMINATION OF PAIRS OF CERAMIC REFERENCE GROUPS★ WITH THE SAME CHEMICAL COMPOSITION★ S for SAME,★ S-S for SAME-SITE, DIF or \X{D}D for DIFFERENT★ T for TYPE, P for PERIOD★ MIX for MIXED★ R-MAT or R-M for RAW-MATERIAL★ CONT for CONTINUITY★ POSS for POSSIBLE, CNTR for center ★ TI for TITLE, LI for LISTSAME SITE, SAME PERIOD, DIFFERENT TYPE ★ (SAME-S-P-DIF-T) ★ S-MIX-R-M-S-TEC ★ (Same mixed raw material and same technology) REGLE : 16-TI-S-MIX-R-M-S-TEC SI NAME PROJECT = VANDALALORS ECRIRENLECRIRE “DIFFERENT CERAMIC WARES MADE FROM THE SAME” ECRIRENL ECRIRE “MIXED RAW MATERIAL AND USING THE SAME TECHNOLOGY” ECRIRENL ECRIRE “DURING THE SAME TIME PERIOD”ECRIRENL ECRIRE “WARE AND WARE SITE PERIOD”ECRIRENLFR FIN ETAPEREGLE : 17-LIST-S-MIX-R-M-S-TECSI SAME-CHEM-COMPOS (A)=(B)SAME-S-P-DIF-T (A)=(B) ALORSSAME-TEC-DIF-TYPE (A) <== (B) S-MIX-R-MAT-DIF-

TYPE (A) <== (B)ECRIRE TYPE(A) " "TYPE(B)" "SITE(a)" "PERIOD(A)ECRIRENLFR FIN ETAPE★ FOR SAME TYPE, SAME PERIOD, DIFFERENT SITE★ (SAME-T-P-DIF-S)★ S-MIX-R-M-FOR-S-T-D-SITE★ (Same mixed raw material for same type at different sites)REGLE : 18-TI-S-MIX-R-M-FOR-S-T-D-SITE SI NAME PROJECT = VANDAL ALORS ECRIRE "SAME CERAMIC WARES MADE WITH THE SAME" ECRIRENL ECRIRE "MIXED RAW MATERIAL AT DIFFERENT LOCATIONS" ECRIRENLECRIRE "DURING THE SAME TIME PERIOD"ECRIRENLECRIRE "WARE PERIOD SITE AND SITE"ECRIRENL FR FIN ETAPEREGLE : 19-S-MIX-R-M-FOR-S-T-D-SITE SI SAME-CHEM-COMPOS (A) = (B)SAME-T-P-DIF-S (A) = (B) ALORSS-MIX-R-M-FOR-S-T-D-SITE (A) <== (B) ECRIRE TYPE(A) " "PERIOD(A)" "SITE(A)" "SITE(B) ECRIRENLFR FIN ETAPE★ POSSIBLE IMPORT FROM A COMMON BUT UNKNOWN SOURCE★ (SAME-T-P-DIF-S continued) ★ (S-MIX-R-M-FOR-S-T-D-SITE continued) REGLE : 20-TITLE-IMPORT SI NAME PROJECT = VANDALALORS ECRIRENLECRIRE "CONTEMPORARY CERAMIC WARES POSSIBLY IMPORTED FROM A" ECRIRENLECRIRE "COMMON UNKNOWN SOURCE"ECRIRENLECRIRENLECRIRE "WARE PERIOD SITE AND SITE"ECRIRENLFR FIN ETAPE.................................REGLE : 22-L-POSS-COMMON-IMPORT SI S-MIX-R-M-FOR-S-T-D-SITE (A) = (B)PERCENT (A) < 10 PERCENT (B) < 10ALORSSOURCE (A)<=== POSSIBLE-IMPORTSOURCE (B) <== POSSIBLE-IMPORT POSS-SAME-IMPORT-CNTR (A) <== (B)ECRIRE "TYPE(A)" "PERIOD(A)" "SITE (A)" "SITE(B)ECRIRENLFR FIN ETAPE

Extracts from VANDAL *printed results*

CERAMIC REFERENCE GROUPS WITH THE SAME CHEMICAL COMPOSITION WARE SITE PERIOD WARE SITE PERIOD BLACK-ON-BF GODIN___ BA4 BLACK-ON-BF GODIN___ BA5 BLACK-ON-BF GODIN___ BA4 UNTEMPERED_ GODIN___ BA4 HARD______ GODIN___ BA4 BLACK-ON-BF GODIN___ BA4 HARD______ GODIN___ BA4 UNTEMPERED_ GODIN___ BA4 UNTEMPERED_ SEH-GABI BA4 UNTEMPERED_ GODIN___ BA4 BLACK-ON-BF SEH-GABI BA4 UNTEMPERED_ SEH-GABI BA4

CONTINUITY IN THE USE OF THE SAME MIXED RAW MATERIAL FOR PRODUCTION OF THE SAME CERAMIC WARES WARE SITE FROM PERIOD TO PERIOD BLACK-ON-BF GODIN___ MC2 BA5 UNTEMPERED_ SEH-GABI MC2 BA4 UNTEMPERED_ SEH-GABI MC2 BA4 HARD______ SEH-GABI MC2 BA4 HARD______ GODIN___ MC3 BA4 BLACK-ON-BF SEH-GABI MC2 MC3

DIFFERENT CERAMIC WARES MADE FROM THE SAME MIXED RAW MATERIAL AND USING THE SAME TECHNOLOGY DURING THE SAME TIME PERIOD WARE AND WARE SITE PERIOD BLACK-ON-BF UNTEMPERED_ GODIN___ BA4 HARD_____ UNTEMPERED_ GODIN___ BA4 HARD_____ BLACK-ON-BF GODIN___ BA4 BLACK-ON-BF UNTEMPERED_ SEH-GABI BA4 HARD_____ UNTEMPERED_ SEH-GABI BA4 HARD_____ BLACK-ON-BF SEH-GABI BA4 BLACK-ON-BF UNTEMPERED_ GODIN__ MC3 HARD_____ BLACK-ON-BF GODIN__ MC3 HARD_____ BLACK-ON-BF GODIN__ MC3 SOFT_____ IMPRESSED_ GODIN__ MC1 ....

CONTEMPORARY CERAMIC WARES POSSIBLY IMPORTED FROM A COMMON, UNKNOWN SOURCE WARE PERIOD SITE AND SITE BLACK-ON-BF MC1 SEH-GABI GODIN___ UNTEMPERED_ MC3 SEH-GABI GODIN___ UNTEMPERED_ BA4 SEH-GABI GODIN___

SAME CERAMIC WARES POSSIBLY TRADED BETWEEN TWO KNOWN SITES WARE ORIGIN IMPORTED TO PERIOD HARD_____ GODIN___ SEH-GABI BA4 HARD_____ GODIN___ SEH-GABI MC3 UNTEMPERED_ GODIN___ SEH-GABI MC2 BLACK-ON-BF SEH-GABI GODIN___ MC2

SAME CERAMIC WARES POSSIBLY CONTINOUSLY IMPORTED FROM THE SAME UNKNOWN SITE WARE SITE WARE SITE FROM PERIOD TO PERIOD UNTEMPERED_ SEH-GABI UNTEMPERED_ GODIN___ MC3 BA4

CERAMIC WARES POSSIBLY CONTINUOUSLY TRADED BETWEEN TWO KNOWN SITES WARE ORIGIN IMPORTED TO FROM PERIOD TO PERIOD HARD___ GODIN__ SEH-GABI MC3 BA4

References

Gardin, J.-C., O.Guillaume, P.-Q.Herman, A.Hesnard, M.-S.Lagrange, M.Renaud, & E.Zadora-Rio 1988. *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology*. Chichester: Ellis Horwood.

Lagrange, M.-S. & M.Renaud 1983. Interpretation des documents figures en archeologie et histoire de l’art: essai de simulation sur ordinateur. In *Image et Céramique Grecques,* Lissarrague, F. & Thélamon, F. (eds), 43–66. Rouen: University of Rouen.

Lagrange, M.-S. & M.Renaud 1985. Intelligent knowledge based systems in archaeology: a computerized simulation of reasoning by means of an expert system. *Computers and the Humanities* **19,** 37–52.

Lagrange, M.-S. & M.Renaud 1988. The SNARK system. In *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology,* Gardin, J.-C., Guillaume, O., Herman, P.-Q., Hesnard, A., Lagrange, M.-S., Renaud, M., & Zadora-Rio, E. (eds), 46–58. Chichester: Ellis Horwood.

Laurière, J.-L. 1986a. *Intelligence Artificielle. Résolution de problèmes par l’homme et la machine.* Paris: Eyrolles.

Laurière, J.-L. 1986b. Un language déclaratif: SNARK. *Technique et Science Informatiques* **3,** 141–171.

Vitali, V. 1989. *Etudes de Provenance Archéometriques des Artifacts Ceramiques. Problèmes actuels de méthodologie.* Research report UPR 315. Paris: Centre Nationale de Recherches Scientifiques.

Vitali, V. & M.-S.Lagrange 1988. VANDAL: an expert system for the provenance determination of archaeological ceramics based on INAA data. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 369–375. Oxford: British Archaeological Reports.

18

Designing a workbench for archaeological argument

ARTHUR STUTT & STEPHEN SHENNAN

Introduction—what the myth of Frankenstein tells us

Societies generate myths of various kinds to regulate the use of powerful devices. Archaeologists are as aware as anyone of the nature of these myths, of their regulatory as well as expressive roles. Often there are two complementary myths; the one celebratory, the other admonitory. For instance, the myth of Pygmalion suggests that, through love, a statue can be brought to life. This myth of control over nature has its contrary in the Frankensteinian myth of destruction in which the creation destroys the creator.

We feel that the notion of an expert system lends itself too readily to the kind of misconception engendered by the myth of control and to the dangers of an uncritical acceptance of the results of such a system. The main problem here is what we call the *oracular expert system*. By this we mean that the system gives a misleadingly authoritative and unquestionable response. At the same time we do not want to deny a place to the technology. We want to temper the myth of control with an awareness that Artificial Intelligence (AI) is a vastly difficult project which may never be fully (or even partially) realized.

Fortunately there is evidence that users resist the deployment of technology unduly influenced by the myth of control. While the earliest expert systems were developed in medicine (Mycin dates from 1976) there is as yet no widespread acceptance of these systems. In archaeology there have been discussions of the possibilities of the technology for nearly 10 years (eg. Ennals & Brough 1982) without any substantial inroads into archaeological practice. This contrasts sharply with the use of expert systems in various areas of business (where presumably profit overrides other concerns) and technology (where the systems are based on clearly definable mathematical theories). Why is this so and what should be done about it?

No doubt the resistance is partially attributable to professional pride in some cases and to sheer stubbornness and reaction in others. However, we suggest that one main obstacle to acceptance is the cognitive dissonance brought about by the discrepancy between what AI systems are claimed to achieve and what they actually achieve; that, in other words, the mistrust of expert systems is perfectly reasonable. As a solution to this problem for potential users in the humanities we suggest that the need is not for oracular expert systems but for tools to support the high level tasks that these researchers and teachers fulfil. We have argued this point in several recent papers (Stutt 1988; Shennan 1990; Stutt 1990) and suggested that, given the role of argumentation in humanities research, what users in the humanities need is a means of supporting argumentation. The task we have set ourselves, then, is not that of reproducing a human arguing agent which we think would be a Frankenstein's monster, if it were possible at all. Rather, we are attempting to produce a system which knows enough about a domain (ie. incorporates knowledge-based technology) and about argumentation to be able to give reasonable criticisms of an argument advanced to it.

This chapter discusses some of the alternatives to the oracular expert system within AI and suggests the reasons why such a system is appropriate for archaeology. It outlines the model of interpretation and argument which underlies our suggested system, puts forward a design for WORSAAE (an argument support program or workbench), and shows the system in action. It is important to point out that this system is in the process of completion and that the results illustrated are not intended as substantive additions to the archaeological knowledge base. The design and its implementation can only be finally assessed in terms of its utility in assisting human arguers to hone their arguments. We thus avoid the trap of claiming that the system can solve all the problems in archaeology, but remain open to the charge of promising more than we can supply.

Some alternatives to the oracular expert system

It is morally objectionable and perhaps dangerous for a system to be presented/present itself as intelligent when it has only limited capabilities. The moral objections which are incurred by any misleading description are here compounded by the moral objections which can be raised by the mere project of constructing artificial human beings. This is further compounded by the moral dilemmas posed by the ease with which advanced computer systems can trade on our innate capacity to endow producers of sentences with intelligence and other human attributes.

There are also political objections. Expertise and authority go hand in hand on local and global scales. Thus the expert (doctor, lawyer etc.) is the one who commands respect and a high salary. On a global scale the expert in a domain commands authority and submission to a particular perspective. It is bad enough that in inter-human transactions the power structures which underlie these are governed by perceptions of relative access to knowledge. That this underlying power structure may be inherited by machine based 'experts' poses political questions of much complexity. For instance is it right that only the wealthy (those with access to the technology) should have access to the knowledge available? This differential exists both within westernized societies and between western and third world societies. It is likely that the differences which already exist here will be exacerbated by the continued development of such systems. We can only guess at the political systems needed to combat the excess of power concentrated in the hands of those with the necessary access.

Critics such as Searle (1980) and the Dreyfuses (1986) have made it clear that there are profound philosophical problems with the notions of an artificial intelligence (see especially Winograd & Flores 1986). The basic problem is that underlying the project of constructing an artificial intelligence there is a particular view of the world— positivistic, formalist, materialist. This view of the world frequently presents itself as the common sense apprehension of reality. However it is obviously as ideologically permeated as any other viewpoint and as subject to the criticisms of other viewpoints.

There are, however, various alternative paradigms for system interaction, which we will briefly outline below (see Stutt 1989 for further details).

The critiquing approach

There are two main bodies of research which exemplify this approach: the work of Langlotz & Shortliffe (1983) on ONCOCIN and of Miller (1983) on ATTENDING, both medical systems. In this approach the program proceeds by asking for a user's solution or other textual construction (eg. for therapy management), compares it with its own and then produces explanations of the differences.

Action-coordination systems

Winograd & Flores (Winograd & Flores 1986; Winograd 1988) argue that since it is impossible to capture all the background knowledge utilized by humans in solving problems it is better to use the computer not as an autonomous problem solver but as a kind of assistant which models only a subset of human skills. They have developed what they call the 'coordinator'. This is basically a system which stands between a group of users and an electronic mail system. It encapsulates knowledge about the structure of conversations and the nature of speech acts; with conversations viewed as a sequence of speech acts. Thus, when a performative such as a promise is made the system acts as an intelligent reminder of the commitment until the promised task is completed.

ARGNOTER

The Colab project as reported by Stefik, Foster, Bobrow, Kahn, Lanning & Suchman (1987) is an attempt to create an environment which can be used to facilitate 'collaborative problem solving'. Users can interact verbally or by means of software which allows them to draw or write on shared screen areas. As a part of this project the team have designed what they call the ARGNOTER which will act as a sort of argumentation spreadsheet. The program provides for the entry of proposals, reasons for and against them and the assumptions behind these reasons. It then requests the various evaluation criteria such as feasibility and cost which the participants want to use and allows the viewing of the proposals in relation to the different criteria.

CRACK, VIEWPOINTS and JANUS

The CRACK constructive design system (Fischer, McCall & Morch 1989) aids the designer in the construction of a design, in this case, of kitchens. The system provides a graphic window with a palette of the different sorts of kitchen furniture. As the designer creates the embryonic design a critic is activated when some conflict between design elements is encountered (for instance, the attempted placement of a refrigerator next to a cooker). The criticism is reported in a critic window and further text-based explanation can be generated.

The VIEWPOINTS system goes beyond the constructive design of the above system to what the authors call argumentative design. VIEWPOINTS acts as a 'look-up manual where designers can find answers to particular problems and consider the various arguments for and against those answers' (Fischer, McCall & Morch 1989, p. 272), and as such is a hypertext system. In addition the user can extend the system by adding his or her own arguments about a particular issue (such as the location of sinks).

The JANUS system (McCall, Fischer & Morch 1990) integrates the graphical construction window and critic of CRACK with the richer argument structures of VIEWPOINTS. Thus, instead of the rule-based criticisms in the critic window, the current state of the argumentation about a particular issue can be displayed if requested.

Constraint based argumentation

The EUCLID system described by Smolensky, Bell, Fox, King & Lewis (1987) uses the structure of reasoned argument as a way of structuring the form which the interaction between system and user takes. To achieve this, the system offers the means for the user to construct arguments. The system has knowledge of the form these must take (captured in a grammar-like Argument Representation Language) and, thus, can be said to understand the structure of an argument while having no understanding of the specific contents of the propositions used. In addition the system offers a set of higher order structures for arguments by analogy, argument refutations and so on. The fact that it is a hypertext system means that more detailed information can be obtained by the user, including the full text of any paper underlying an argument. The system provides many facilities for the generation of arguments (by the user). These include templates based on the high-level structures, aid in inserting structuring relations and checks for missing elements of these templates. While this structuring aid is felt to be effective, the authors go on to assert in an important aside (Smolensky, Bell, Fox, King, & Lewis 1987, p. 232) that they ‘feel that the process of expressing an argument in ARL form requires extensive human processing and that the prospects for automating the process are dim’.

Our alternative—WORSAAE

WORSAAE is an acronym for a WORkbench Supporting Archaeological Argument Exploration. The WORSAAE project is aimed at producing a program which can be used by the archaeologist as an aid to the production and evaluation of arguments. This program is currently under construction at Southampton University. When the system is completed it will allow the user to:

1. display arguments.
2. manipulate arguments.
3. request arguments embodying site interpretations.
4. view these as text and/or graphics where appropriate.
5. enter and store arguments.
6. have these criticised by the system.
7. have alternative arguments identified.
8. have arguments evaluated.

This approach combines both the explicit adviser and the graphic manipulation approaches identified by Woods (1986), and does this by combining techniques from hypertext (see Conklin 1987) and AI. This combination has been suggested by many researchers (eg. Halasz 1988) but WORSAAE is one of the first systems to try to implement these ideas. Halasz (1988, p. 847) suggests: ‘The integration of hypermedia and AI technology is an interesting direction to explore. In many ways, hypermedia and knowledge-based systems are a natural fit.’ Argumentation (as the EUCLID researchers also believe) is a natural means of integrating these two technologies, since this integration is best achieved not at the implementational level but in terms of a model of a higher level cognitive process which necessitates the combination of the discursiveness of hypertext with the inferential capabilities of expert systems. For our purposes this cognitive process should also be an essential aspect of research in the humanities. We suggest that a computational model of argument fits this description. Knowledge-based inference is necessary for argument generation and assessment and for selecting an appropriate response. Hypertext is necessary for the storing and display of an argument. The resources of a hypertext document could also be used for exemplification and illustration. The resulting system is one which has much in common with human/human arguments but which differs in significant ways. It can display premises for an argument at the same time in different windows as both text and graphics. It can display all the links between claims and grounds and between supporting and attacking arguments and can provide the means for navigating around these links. Most importantly, however, this system does not pretend to act as an ‘arguer’ but is transparently a machine oriented means of analysis and assessment.

The particular borrowings from the systems discussed above which have found their way into WORSAAE include: the notion of critiquing from ATTENDING and CRACK; the notion of differential analysis from ATTENDING; the use of a model of the discourse structure from the COORDINATOR; particular forms of argument layout and aid in entering these from ARGNOTER, Marshall’s work and EUCLID; the notion of an active critic from CRACK; the idea of an argument grammar from the ARL of EUCLID; the combination of argument and hypertext from Marshall, VIEWPOINTS and EUCLID.

Archaeological needs

We believe that WORSAAE will satisfy the demands of a variety of recent contributors to the ongoing discussion about the nature of archaeological inference. The various knowledge bases will include in an explicit form the heuristics needed for the interpretation of site data and will thus function as catalogues of justificatory propositions in the manner suggested by Schiffer (1976). At the same time these rules can be debugged and made consistent since they are part of a runnable computer program. The system's knowledge bases also act as a version of the sets of transformations studied by Gardin and his colleagues (Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988). WORSAAE can be seen as a part of the production of hypertext-like reports suggested by Rahtz, Carr & Hall (1989). Finally, the system as a whole satisfies the recent demands made by Hodder (1989) and Tilley (1989) for a new form of writing in archaeology in which the dialogic nature of the process of excavation and report writing is captured.

> Few sites can be dug in total consensus. The disagreements should spill over into the text so that the reader can insert herself into a process of argument rather than having to consume pre-packaged, supposedly neutral fare. (Hodder 1989, p. 273)

We suggest that a mixed hypertext/AI system for the support of archaeological argumentation best fulfills all these demands while at the same time acting as a realistic test of the claims of the utility of such systems to archaeologists.

The model of argument

The term 'argument' is ambiguous between a set of structured propositions and a conversational exchange. The two senses are, of course, not unrelated—the conversational exchange can be borne by a series of structured sets of propositions and can result in them. However, concentration on the notion of arguments as structured sets of propositions, especially by logicians, has made it difficult to produce a credible model of the arguing process. Current work on the rhetoric of argumentation (Rowland 1987) offers a corrective here. Following the seminal work on the distinction between the two senses of argument by O'Keefe (1977) we view argument as primarily interactionist or dialogic in nature, only making use of structured sets of propositions as a means to an end. As Blair & Johnson (1987, p. 51) say:

> …argumentation is an intricate, many levelled interchange of pro and con considerations, not a one-shot demonstration that settles the question once and for all.

According to our model there are three levels at which argument can be discussed:

(a) Argument-2. This can also be referred to as debate—the statement of a topic or issue followed by a series of arguments for and against. In computational terms it involves the representation and manipulation of a complex data structure which records the arguments for and against particular propositions.

(b) Argument-1. Each argument for or against a proposition in a debate can be viewed as itself having a structure. This is most commonly represented as a tree-like structure in which propositions at the root node are supported by arguments and sub-arguments represented by nodes at the various lower levels. The outlining facilities provided by modern word-processing software use indentation to capture these structures (and can thus be regarded as primitive argument support programs).

(c) Argument-0. Each argument (or micro-argument) is made up of a single inference step. Following the work of Toulmin (1958; Toulmin, Rieke & Janik 1979) this can be viewed as a complex data structure in which a claim is related to the evidence (or data) it is based on by means of a warrant (a piece of knowledge allowing the derivation of the claim from the evidence). The warrant can be of many different types corresponding to the mode of inference used; these include deduction, analogizing and causal reasoning. The warrant in turn is grounded in a backing drawn from theoretical or common sense knowledge. Toulmin and his colleagues illustrate these notions by an example drawn from palaeontology (Toulmin, Rieke & Janik 1979, p. 253):

Claim: The earliest known anthropoid apes lived in the African Rift Valley

Ground: Geological and palaeontological reports from Africa, China, Java, etc.

Backing: Our experience in developing a systematic interpretation of palaeontological evidence indicates that

Warrant: The presence of fossilized anthropoid remains in earlier rock formations in one place than another indicates the earlier existence of living anthropoids in one place than another.

Warrants can be modelled as the If-Then heuristics of expert systems technology. Backing theories can be modelled as sets of facts, assumptions and rules which act not only as a way of grounding arguments but constrain the selection of lower level rules in the inference processes which produce arguments. These latter aspects ensure that the representation of an argument is multi-dimensional in that an argument-2 can encompass not only many different conclusions and sequences of justifications but also many different reasoning modes and theoretical viewpoints.

The model outlined above can be used as the basis for the introduction of a semi-formal language for the representation of arguments (which has similarities to the ARL of Smolensky, Bell, Fox, King, & Lewis 1987). We have presented this as a series of productions (or rules governing possible argument patterns):

argument–1 :== claim | claim *because* justification | justification therefore claim | justificationclaim :== a proposition justification :== ground (*and* ground) ⋆ ground :== a proposition | argument-1

In the above the symbol | represents alternatives, ⋆ represents possible repetition and () optional elements. The italicised elements represent terminal symbols (the elements of the language being defined as opposed to the symbols required by the grammar). Thus, for example, the first production asserts that an argument-1 can be made up of a claim or a claim followed by the word ‘because’ followed by a justification or a justification followed by the word ‘therefore’ followed by a claim or a justification (with the claim implicit).

The above grammar captures several important features of everyday (and indeed scientific) argumentation (cf. Cohen 1987): first, it has recursive elements— arguments are often made up of sub-arguments and naturally form hierarchically organized trees; second, it is frequently gappy; and, third, it can be expressed in more than one way (more indeed than the above grammar captures). Often an argument consists only of a claim (even in extreme cases solely of a ground). We have to provide the inferences necessary to fill these gaps. It is one important consequence of drawing attention to the structure of argumentation that it makes it obvious when there are gaps which require filling.

The above grammar has its limitations. It captures in a succinct form the essentials of the makeup of argument-1s. However, grammars of this kind are inappropriate for representing the dialogic interchange which creates argument-2 structures. As we have seen, an argument-2 can be viewed as a network of related argument-1s where the relations are, for instance, *supports, attacks, restates*.

In what follows we describe what is involved in the conducting of an overall argument-2 in terms of a set of strategies which operate at a high level to produce argument structures. Although the literature on the representation of the rules covering such dialogic exchanges is vast (Reichman 1985; Winograd & Flores 1986; Baker 1988) much of this is concerned with everyday conversational exchanges or with tutorial interactions. We will therefore confine ourselves to a discussion of the strategy rules for the conduct of an academic debate. In any debate (inside archaeology or without) there will be at least two and possibly several argument-1s competing from different points of view. These arguments will be related to others by relations of support and attack, restatement and so on. At the same time, since the debate has an overall purpose and goal, these arguments will not be randomly produced but will be generated according to rules which we might call strategic. Frankel (1987) has codified our intuitive conception of the strategies involved in scientific (knowledge-producing or epistemic) debate. A debate, while dialogic in character, is more formal than everyday conversation and hence not involved in many of the issues involved in spontaneous argumentational exchanges. Frankel evolved four main strategies for the conducting of an academic or epistemic debate.

Strategy S1: Theory T1 leads to a solution/interpretation, which deals with some aspect hitherto undealt with, allowing the claim that T1 is more likely.

Strategy S2: In the face of this solution/interpretation proponents of T2 should either:

Strategy S2a: find another solution/interpretation, or
Strategy S2b: attack that of their opponents.

Strategy S3: In the face of successful S2 moves proponents of T1 can either:

Strategy S3a: attack the T2 solution, or:
Strategy S3b: defend their own.

They can defend their own by

Strategy S3b1: changing it;
Strategy S3b2: getting a new one; or,
Strategy S3b3: attacking the attack on it.

Strategy S4: In the face of an interpretation/solution with two independent lines of support, opponents should: (a) attack it and (b) look for an alternative.

DOMAIN INDEPENDENT	ACROSS SEVERAL DOMAINS	DOMAIN DEPENDENT
interpersonal knowledge		
logical rules e.g transitivity		
semantic rules e.g. what 'red' means	semantic rules e.g. what 'section' means	semantic rules e.g. what 'Harris matrix' means
heuristics—everyday 'rules of thumb'	heuristics e.g. that lower in a section=older	heuristics e.g. 'Worsaae's Law'
strategies for reasoning e.g. hypothesize and validate	strategies for reasoning e.g. statistical inferences	strategies for reasoning e.g. reasoning from ethnographic analogies
strategies for debate e.g. undermine and replace	strategies for debate e.g. rules for plate tectonic debate	strategies for debate e.g. find better analogies?? ascend to epistemology??
common sense—factual	chemical facts e.g. iron rusts	facts about specific sites e.g. St Veit-Klinglberg
common sense—theory e.g. how thermostats work	general theory e.g. evolution/ecology or marxism	specific theory e.g. Schifferian transformation theory

Figure 18.1 The sorts of knowledge involved in archaeological argumentation.

In a recent paper (Stutt & Shennan 1990) we discuss this work in more detail and show how it can be applied to two examples from archaeology: (a) The debate about the nature of the Star Carr site and (b) the debate about the origin of social stratification in bronze age Europe.

Finally, we include in our model the notion that argument is knowledge-intensive. In comprehending an argument (and we must comprehend it to assess it) we make use of various kinds of knowledge including common sense knowledge and logic. However, we will only be able to fully grasp the import of an argument if we have a deep knowledge of the domain and grasp fully its current pre-occupations. This is a point frequently overlooked (or hidden) in positivistic accounts of scientific arguments. Semantic and, indeed, pragmatic considerations are involved, as well as syntactic, in producing or understanding an argument.

> …in general if you are to assess scientific arguments in their full complexity, you need to know their history and current scientific opinion —to know, so to speak, the rules of the game. (Fisher 1988, p. 12)

We have summarized the forms of knowledge applicable to argumentation in archaeology in Fig. 18.1.

A model of interpretation

A model of interpretation is needed as part of the overall model of argumentation since we need to specify the component which produces arguments or which is used for argument criticism. In the domain of archaeology the principal reasoning task is interpretation and the principal object of argumentation the results of interpretation. Other tasks such as cataloguing, description and statistical analysis arguably either contain interpretive elements or can be seen as part of the interpretive process. While the model of interpretation forms part of our overall model of argumentation it can also stand on its own as a model of the interpretive task. For this reason we have discussed it in a separate section.

-----**harder for lay person to understand**-------->

We have derived the model of interpretative reasoning from the work of J.-C. Gardin (Gardin 1980; Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988) and various of his colleagues, in which he puts forward what he calls a *logicist analysis* of archaeological reasoning. Gardin's logicist models provide a clear, principled and general statement of interpretive reasoning. They are derived from a consideration of archaeological reasoning and therefore are ideal for the model of interpretation needed for a model of argumentation in that domain. The main aim is to produce an architecture of the propositions which make up an interpretation. There are three types of proposition: (a) those which capture the initial data forming the basis for an interpretation, or P_o propositions; (b) those which capture the end product of the process of reasoning involved in the interpretation (P_n); and, (c) the intermediate propositions which capture the stages in the reasoning intermediate between initial and terminal propositions (Pi). A tree-like representation of an interpretation can be produced in which the connections between nodes are determined by the application of transformation rules. The type of the resulting relation between the nodes depends on the type of the transformation. Possible transformations include various forms of reformulation and of inference. For instance, a transformation may be applied to a node to derive a deductively valid proposition which can serve as a node at a higher level; the relation being one of implication. The sorts of transformations which are operative can be expressed as rules in a production system. Gardin et al. (1988) and Lagrange & Renaud (1985) describe the results of experiments in the use of production rules to capture the transformations.

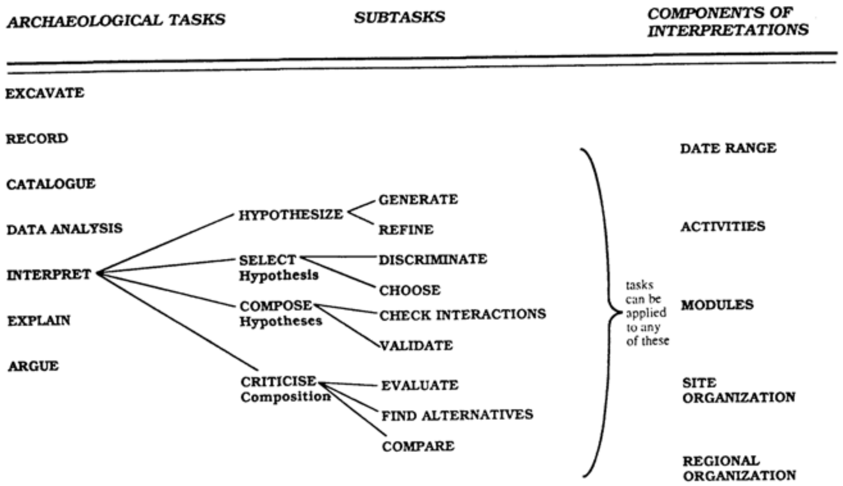


Figure 18.2 A task analysis for archaeology.

In an earlier application of this model, archaeological interpretation was taken to have the following components:

1. Classification of features and finds. The term ‘features’ encompasses the various aspects of archaeological sites both man-made and natural (eg. pits, ditches, walls etc.). ‘Finds’ are mobile features which are either man-made or natural such as bone fragments.
2. The reconstruction of past human activities in terms of activity areas and their associated activity. An ‘activity area’ is a significant area of a site at which identifiable human activities (eg. cooking or hide-working) were carried out.
3. Cultural interpretation. That is, the creation of an interpretation or cultural profile for the site as a whole which includes a determination of the technology, subsistence, social organization and religious or other beliefs of the occupants of the site.

Due to certain difficulties with the model as presented above (especially in its lack of a top-down element) we evolved an expectation-based model of archaeological reasoning. If we incorporate this top-down element we obtain a model of interpretation in which propositions are produced initially either by applying rules in a bottom-up manner or by making use of an ideal description of a site. Each node may then be subject to several cycles of bottom-up reasoning and top-down checking before the final value is arrived at. This may involve a reappraisal of all the lower nodes in the tree which constitutes the partial interpretation. Thus, by gradually extending the satisfactory nodes the final interpretation is produced. This expectation-based model of interpretation has been implemented for the domain of archaeology and reported in Patel & Stutt (1988). The system we produced was called KIVA and reasoned about the domain of Pueblo Indians from the south-west of the USA, a kiva being a ceremonial area in a pueblo village. Typically a series of interpretations would be produced which could be pruned subsequently by the application of constraint rules embodying the model of an idealized pueblo site.

The above model has been extended more recently to include an analysis of the tasks and sub-tasks which make up the interpretation process. This has come about as a result of knowledge acquisition sessions between the authors (see Fig. 18.2). This model seems to have certain advantages over those already discussed in the archaeological literature in that it allows the construction and partitioning of a usable knowledge base and the resultant efficient firing of rules. The extensions mean that the model includes not only knowledge about a domain (the facts and rules which relate these) but also knowledge of the order in which these are applied (reasoning strategies as opposed to debate strategies—see Fig. 18.1), as well as the different tasks which make up the process of interpretive reasoning. The domain model thus includes elements of strategic and task knowledge as well as factual and relational knowledge.

In short, the model of archaeological reasoning we have adopted views it as composed of a series of tasks including classification, interpretation and explanation. Interpretation may be regarded as part of the explanation task and is further decomposable into a cycle of hypothesis, selection, criticism and composition. The process of archaeological interpretation is thus viewed as a multi-stage cycle with top-down and bottom-up elements, in which the reasoning is plausible (frequently analogical) and governed by theoretical principles.

The components shown in Figure 18.2 represent the sorts of topics which archaeological interpretations may focus on. The range of sub-tasks can be applied in establishing values for any of these components.

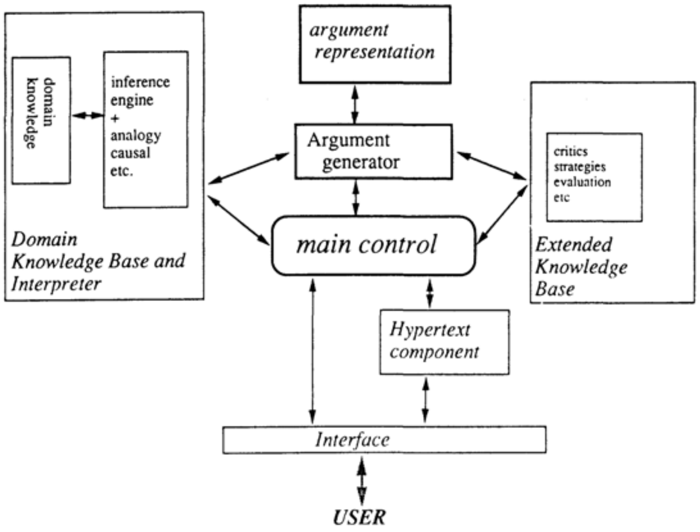


Figure 18.3 WORSAAE.

WORSAAE

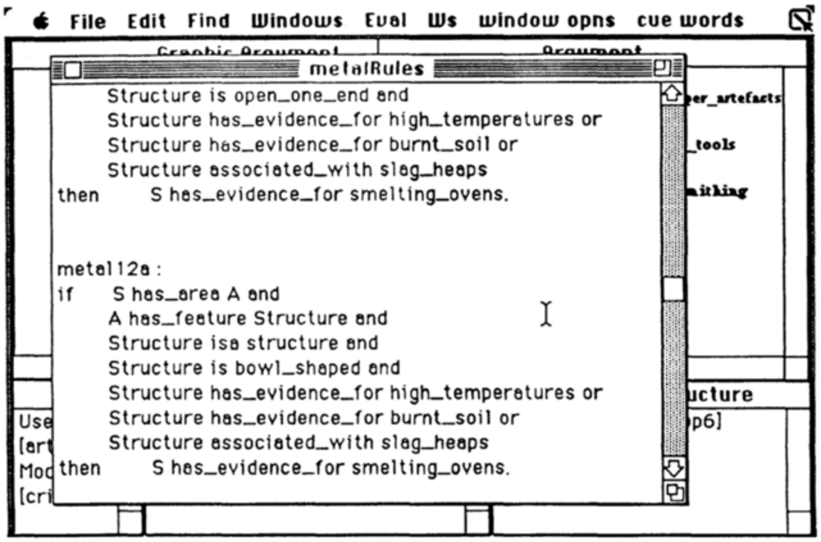
WORSAAE—the knowledge bases

As can be seen from Figure 18.3, the WORSAAE system is relatively complex. However in some ways this is deceptive. The system is basically a standard knowledge-based system (The Domain Knowledge Base and Interpreter) overlain by various modules for the manipulation and display of the results of this interpreter. We do not propose to discuss the workings of the knowledge base interpreter since these have been discussed frequently in the literature (Ennals & Brough 1982; Bishop & Thomas 1984; Huggett 1985; Lagrange & Renaud 1985; Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988; Patel & Stutt 1988; Patel & Stutt 1989; Francfort 1990). At the time of writing (late 1990), the components implemented include the basic interpreter, the argument representation module, a window, mouse and menu interface and part of the system knowledge base. We will confine our discussion to the knowledge bases since these capture the domain and strategic knowledge necessary for argumentation, as we have seen above, as well as the rules for criticism and evaluation.

Domain An argument in any domain requires the deployment of considerable amounts of domain knowledge. The domain knowledge base fulfills this role by capturing facts and heuristics relating to the St Veit-Klinglberg site (Shennan 1989) which has been interpreted as an early bronze age possibly seasonal settlement of copper producers. The heuristics which have been included in the knowledge base therefore concentrate on the knowledge needed to interpret such data. It should be noted that while we have concentrated on this site and the heuristics needed to interpret it, there is no sense in which the system as a whole is tied to this particular knowledge base. In principle, it should be possible to slot in a variety of knowledge bases. We have concentrated on one site since we feel that no previous system deals with the whole of a complex, real and substantial reasoning task. We have, therefore, attempted to produce a system which can support the complete interpretation of this early bronze age site in Austria, and which, in doing so, solves a realistic problem. Figure 18.4 illustrates the form and content of a typical rule.

System The system knowledge bases are at the heart of WORSAAE's capabilities. Insofar as the system captures the structure of the argument as it is constructed over time in user-computer exchanges, the most fundamental of these is the argument representation module. As a store of the current state of the current argument-2, this is made up of a network of interconnecting nodes, each of which represents an argument-1 which can be decomposed into argument-0s. Each node in the argument network is stored as a frame-like structure of slots and values with links to other nodes and means for traversing these. This common AI representation easily maps onto a hypertext system of nodes and links (see Stutt 1989 for further details of this sort of system). The system also includes a knowledge base which captures rules for the critical assessment of user arguments, and work is proceeding to include the strategic rules discussed above as well as the rules for evaluating arguments. In order to illustrate these we present some examples from the knowledge base of rules for criticism.

The rules for criticism (critics for short) operate at various levels ranging from general through general archaeological and specific archaeological to strategic. The general critics include rules which search for contradiction and self-contradiction as well as more general semantic errors. Some of these can be viewed as the equivalent of the traditional fallacies (see Hamblin 1970; Walton 1989). The general archaeological rules (which can be said to capture some of the semantics of the domain of



```
      Structure is open_one_end and
      Structure has_evidence_for high_temperatures or
      Structure has_evidence_for burnt_soil or
      Structure associated_with slag_heaps
then      S has_evidence_for smelting_ovens.


metal12a :
if     S has_area A and
      A has_feature Structure and
      Structure isa structure and
      Structure is bowl_shaped and
      Structure has_evidence_for high_temperatures or
      Structure has_evidence_for burnt_soil or
      Structure associated_with slag_heaps
then      S has_evidence_for smelting_ovens.
```

Figure 18.4 The form and content of a typical rule.

archaeology) search for inconsistencies between propositions entered and generally held archaeological notions, such as Binford's views on site structure (see below). The specific archaeological critics look for incompatibilities between user argument propositions and the specific knowledge stored in the domain knowledge base(s). Strategic critics (not illustrated) search for inadequacies in the strategy employed by the user.

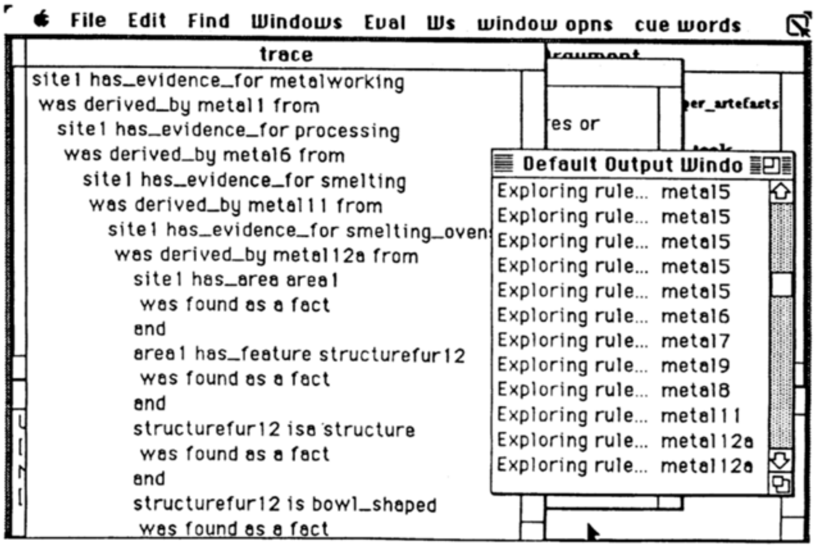
The general critics are activated when a contradictory duo of propositions are input (though not necessarily in the same session). These critics are of three kinds: the generallogical-inconsistency critic responds to a contradiction produced by two different values for the same attribute; the general_self_inconsistency critic responds to an inconsistency between a proposition in the current argument and another used by the same user in a previous argument; the general_semantic_inconsistency critic is activated in the more interesting case where the contradiction is produced when the two values for the attribute are found by a process of inference to be semantically incompatible (see example below). For the above purposes the domain concepts known to the system are divided into those which can have more than one value and those which can have only one value. In the former category lie such concepts as has-find, has_feature, contains, and so on. In the latter category there are concepts such as has_location and has_date. We illustrate the rules for these critics with the following (English) version of a critic:

general_logical_inconsistency if
there are two propositions P1 and P2 and
P1 is Subject Attribute Value1 and
P2 is Subject Attribute Value2 and
it is not the case that Value1=Value2.

The above critics illustrate the use of general logical and semantic knowledge. In contrast, the arch_semantic_inconsistency critic makes use of general archaeological knowledge (see example below). Here the knowledge is based on a set of constraint rules of the form A, B, C where A might be sleeping, B eating and C butchery. The rule suggests that it is unlikely that we will discover sleeping and eating and butchery together. These rules are based on a mixture of our own intuitions and Binford's discussion of the sorts of rules which govern site development (eg. Binford 1983, p. 170: 'It is generally true, in fact, that activities like roasting and butchery, which monopolize considerable amounts of space, are located away from areas used intensively on a day-to-day basis.').

Other critics utilize a specific rule-set such as that for the interpretation of metal processing. The arch_factual_error critic reports on factual inaccuracies. The arch_derivation_error critic is activated when the system is unable to find an input proposition as a fact or to derive it from the knowledge base of facts and rules. This is distinguished from the next critic in that the system ignores the other propositions input by the user in attempting its inferences.

The simple_structure_failure critic (see example below) is activated when the system is unable to derive a proposition representing a claim from propositions representing grounds as the user imposed argument structure suggests should be possible. Unlike the previous critic the system tries to derive the claim from the grounds as specified by the user. In the example below this entails searching for a rule in the domain knowledge base which could allow the derivation. In future



```
site1 has_evidence_for metalworking
 was derived_by metal1 from
   site1 has_evidence_for processing
    was derived_by metal6 from
      site1 has_evidence_for smelting
       was derived_by metal11 from
         site1 has_evidence_for smelting_oven
          was derived_by metal12a from
            site1 has_area area1
             was found as a fact
            and
            area1 has_feature structurefur12
             was found as a fact
            and
            structurefur12 isa structure
             was found as a fact
            and
            structurefur12 is bowl_shaped
             was found as a fact
```

```
Exploring rule...  metal5
Exploring rule...  metal5
Exploring rule...  metal5
Exploring rule...  metal5
Exploring rule...  metal5
Exploring rule...  metal6
Exploring rule...  metal7
Exploring rule...  metal9
Exploring rule...  metal8
Exploring rule...  metal11
Exploring rule...  metal12a
Exploring rule...  metal12a
```

Figure 18.5 An example of the result of an interpreter run.

versions it will be possible to search for a chain of inferences in which the grounds figure as premises at some point with the claim as the final outcome.

The differential-analysis critic displays the result of a differential analysis of the user argument as specified by the argument structure and an argument found by the system, leaving the judgment as to relative merits to the user. This is potentially the most time-consuming criticism sub-mode and is similar to the procedure of the ATTENDING system discussed above.

WORSAAE operation

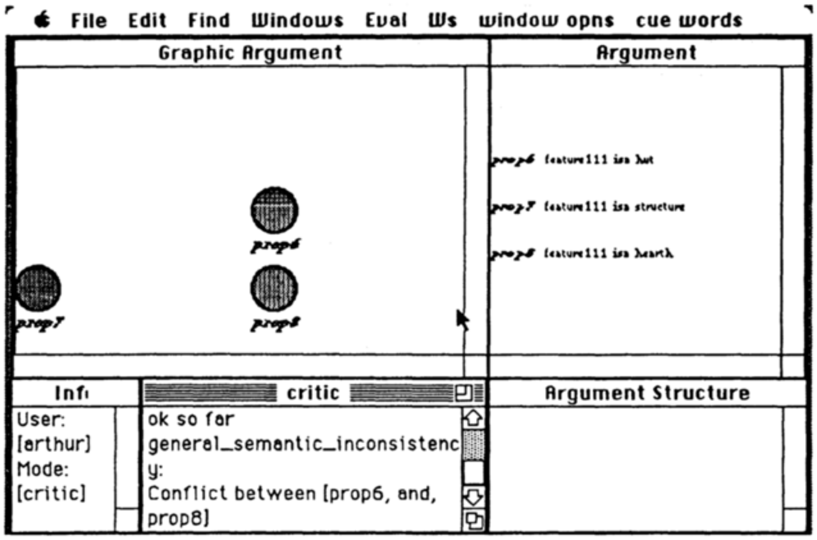
WORSAAE has the following modes: knowledge base enquiry, hypertext viewer, argument viewer, critic. Since all of these are made use of in the critic mode we will concentrate on that.

In essence the system operates as follows. The user either types an argument into the Argument Dialogue window or constructs one using the tools provided in the Graphic Window. The relationships between the propositions of the argument are entered in the graphic window by inserting linking lines of various types. In the text window a menu of cue words can be used to insert the necessary links and give structure to the argument; the separation of propositions from structure foregrounds the function of this structure. The cue words (‘because’, ‘since’) are derived from the grammar given above and represent two possible orderings for an argument structure. When a selection is made from this menu the user is given a further menu of all the possible propositions so far entered. The user chooses the claim and grounds of the argument from these. The structure of the argument thus elicited is displayed in the Argument Structure window. When an argument proposition is entered the system automatically checks the proposition to see if there are any criticisms, thus bringing into play the knowledge base of argument critics. If there are, this is reported in the critic window. The user continues to input propositions and select levels of criticism until s/he is satisfied. At this point the user can have the structure stored for future emendation.

In future systems, tools will be provided to link together the argument-1s constructed as above. It will then be possible to browse argument-2s (system or user originated) using the hypertext component. The user will be able to navigate through the chain of argument-0s which make up the argument-1s which in turn compose the current argument-2. The micro-arguments will be displayed as what we call Toulmin Structures (after Toulmin 1958, cf. Marshall 1987). The evidence which makes up the grounds for the finally grounded argument-0s will be displayed in textual and/or graphical form as appropriate. As with the system discussed above (Smolensky, Bell, Fox, King, & Lewis 1987) the discourse structure provides the necessary navigational structure for the user.

The system in action

The examples of the operation of WORSAAE which follow depend on the WORSAAE knowledge base for St. Veit-Klinglberg. At the moment the user has to enter argument propositions in a constrained manner utilizing, in effect, the syntax of the underlying knowledge base interpreter. In future, it may be possible to extend the system to cope with some natural language processing. Until then each proposition must be a triple of subject, attribute and value {S, A, V}. Subjects are individuals from the sets: sites, site-areas, features, finds and so on. Attributes are descriptional categories fitting each subject category. Thus sites can have the attribute has_date but not the attribute has_colour which might be more fitting for a find.

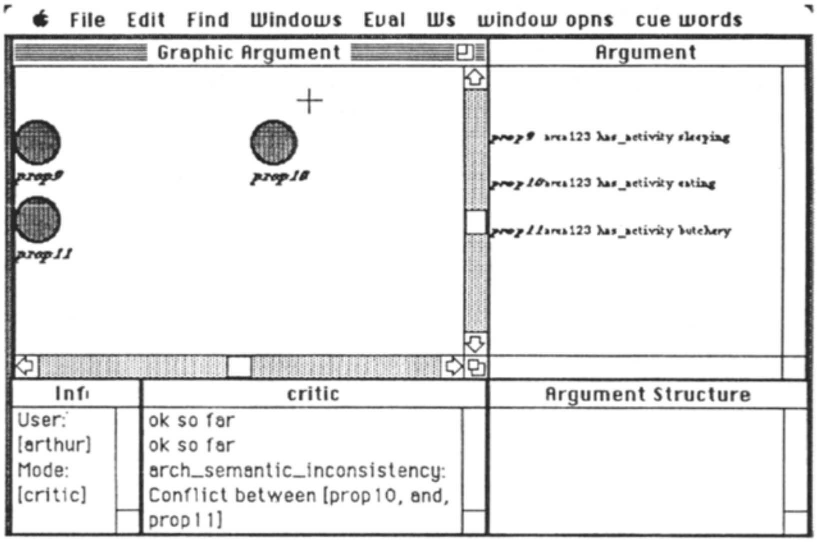


prop6 feature111 isa hut
prop7 feature111 isa structure
prop8 feature111 isa hearth

User:
[arthur]
Mode:
[critic]

ok so far
general_semantic_inconsistency:
Conflict between [prop6, and, prop8]

Figure 18.6 An example of the use of the general_semantic_inconsistency critic.



prop9 area123 has_activity sleeping
prop10 area123 has_activity eating
prop11 area123 has_activity butchery

User:
[arthur]
Mode:
[critic]

ok so far
ok so far
arch_semantic_inconsistency:
Conflict between [prop10, and, prop11]

Figure 18.7 An example of the operation of the arch_semantic_inconsistency critic.

The principal attributes are: has_area, has_artefact, has_find, has_location, is, isa, has_date, made_at, has_module, has_activity. A more general attribute (applicable to all categories of subject) is has_evidence_for. The values of these attributes are those applicable to the categories of attribute; for example, x, y pairs for locations.

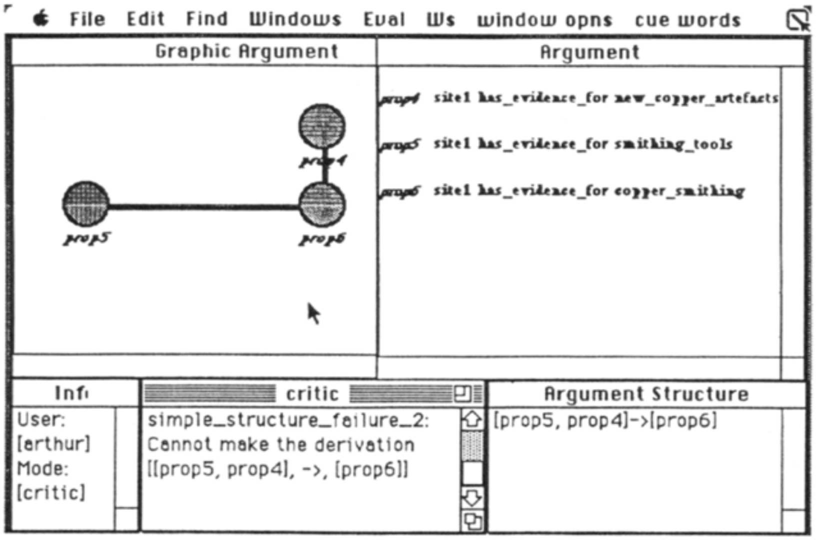
In Figures 18.4 and 18.5 we see the system operating as a rule interpreter (interpreter or expert system). The rule interpreter is based on that given in Bratko (1986). Figure 18.4 illustrates part of the knowledge base for the metal processing sub-component of the domain knowledge base. Figure 18.5 illustrates a trace of the result of an interpreter run showing the rules attempted and the chains of inferences made.

In Figures 18.6–18.8 we see the critics at work. Figure 18.6 illustrates the general-semantic_inconsistency critic. The system reports that feature 111 cannot be both a hut and a hearth. It has no problems with it being both a hut and a structure.

Figure 18.7 illustrates the arch_semantic_inconsistency critic. In this example the system makes use of the general archaeological rule which suggests that eating and butchery cannot take place in the same place. Sleeping and eating are possible (indeed likely) collocated activities.

While the previous two examples illustrate examples where the system determines that it is inconsistent to hold two propositions, Figure 18.8 shows how user determined argument structures may fail. This is illustrated using the simple-structure failure critic. In this example the user has input three propositions about smithing:

- site 1 has evidence for new copper artefacts
- site 1 has evidence for smithing tools



prop4 site1 has_evidence_for new_copper_artefacts

prop5 site1 has_evidence_for smithing_tools

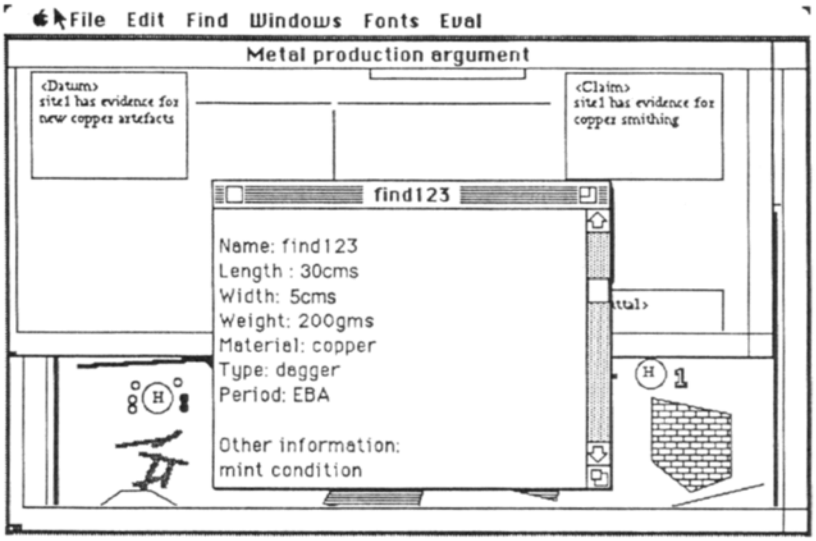
prop6 site1 has_evidence_for copper_smithing

User:
[arthur]
Mode:
[critic]

simple_structure_failure_2:
Cannot make the derivation
[[prop5, prop4], ->, [prop6]]

[prop5, prop4]->[prop6]

Figure 18.8 An example of the simple_structure_failure critic.



<Datum>
site1 has evidence for new copper artefacts

<Claim>
site1 has evidence for copper smithing

Name: find123
Length : 30cms
Width: 5cms
Weight: 200gms
Material: copper
Type: dagger
Period: EBA

Other information:
mint condition

Figure 18.9 An illustration of the hypertext aspect of WORSAAE.

- site 1 has evidence for copper smithing

and suggested that the final proposition can be derived from the previous two. The system replies with the result of the simple_structure_failure critic suggesting that there is no rule which allows the derivation as suggested by the user.

In Figure 18.9 we see a hypothetical example showing the hypertext component. This example shows the Toulmin Structure for an argument similar to the above for copper smithing. This structure can be browsed in many directions; from claim to the succeeding micro-argument, from ground to the preceding micro-argument. The Toulmin Structure also allows the browsing of warrants and backing theories (if these are available) and of any further graphical or textual information relevant to the datum (or ground).

Conclusion

We view argumentation as a process based on producing a complex, multi-dimensional structure which underlies possible representations in natural language and can be captured in a semi-formal language. Archaeological argumentation can be distinguished from argumentation in other domains not only by the specific domain knowledge utilized but also by the inference modes commonly used, the types of theoretical backing given to warrants and the strategies employed. WORSAAE is intended to give the archaeologist support in producing arguments. It does this by utilizing various knowledge bases of general (logical, semantic and strategic) knowledge as well as domain knowledge (both general and specific) to aid the building of argumentational structures in the work-spaces provided. While the system remains only partially realized as yet, our model of argumentation allows the implementation of a system which can utilize knowledge based technology in archaeology without imposing possibly inadequate interpretational models on the user; thus acting as an assistant rather than overlord.

Acknowledgements

The WORSAAE project is funded by grant no. GR/F43222 from the Science-based Archaeology Committee of the UK Science and Engineering Research Council, to whom we are extremely grateful. We are also grateful for the cooperation and help we have received from the Human Cognition Research Laboratory at the Open University, Milton Keynes, UK.

References

Baker, M. 1988. *Arguing with the Tutor: a model for a tutorial dialogue in uncertain knowledge domains.* Milton Keynes: The Open University.

Binford, L.R. 1983. *In Pursuit of the Past: decoding the archaeological record.* London: Thames & Hudson.

Bishop, M.C. & J.Thomas 1984. 'BEAKER'—An expert system for the BBC Micro. In *Computer applications in archaeology 1984,* 56–62. Birmingham: University of Birmingham.

Blair, J.A. & R.H.Johnson 1987. Argumentation as dialectical. *Argumentation* **1(1),** 41–56.

Bratko, I. 1986. *Prolog Programming for Artificial Intelligence.* London: Addison-Wesley.

Cohen, R. 1987. Analyzing the structure of argumentative discourse. *Computational Linguistics* **13(1–2),** 11–24.

Conklin, J. 1987. Hypertext: a survey and introduction. *IEEE Computer* **20(9),** 17–41.

Dreyfus, H.L. & S.E.Dreyfus 1986. *Mind Over Machine.* Oxford: Basil Blackwell.

Ennals, J.R. & D.R.Brough 1982. Representing the knowledge of the expert archaeologist. In *Computer Applications in Archaeology 1982,* 133–44. Birmingham: University of Birmingham.

Ennals, J.R. & J.-C.Gardin (eds) 1990. *Interpretation in the Humanities: Perspectives from Artificial Intelligence.* Library and Information Research Report 71. London: British Library.

Fischer, G., R.McCall & A.Morch 1989. Design environments for constructive and argumentative design. In *Proceedings of the ACM conference on Human Factors in Computing Systems (CHI-89),* 269–275. New York: Association of Computing Machinery.

Fisher, A. 1988. *The Logic of Real Arguments.* Cambridge: Cambridge University Press.

Francfort, H.-P. 1990. Modelling interpretative reasoning in archaeology with the aid of expert systems: consequences of a critique of the foundation of inferences. In *Interpretation in the Humanities: Perspectives from Artificial Intelligence,* Ennals, J.R. & J.-C.Gardin (eds), 101–129. London: British Library.

Frankel, H. 1987. The continental drift debate. In *Scientific Controversies,* Engelhardt, H.T. & Caplan, A.L. (eds), 203–248. Cambridge: Cambridge University Press.

Gardin, J.-C. 1980. *Archaeological Constructs.* Cambridge: Cambridge University Press.

Gardin, J.-C., O.Guillaume, P.-Q.Herman, A.Hesnard, M.-S.Lagrange, M.Renaud, & E.Zadora-Rio 1988. *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology.* Chichester: Ellis Horwood.

Halasz, F.G. 1988. Reflections on notecards: seven issues for the next generation of hypermedia systems. *Communications of the ACM* **31 (7),** 836–852.

Hamblin, C.L. 1970. *Fallacies.* London: Methuen.

Hodder, I. 1989. Writing archaeology: site reports in context. *Antiquity* **63(239),** 268–274.

Huggett, J. 1985. Expert Systems in Archaeology. In *Current Issues in Archaeological Computing,* Cooper, M.A. & Richards, J.D. (eds), 123–141. Oxford: British Archaeological Reports.

Lagrange, M.-S. & M.Renaud 1985. Intelligent knowledge based systems in archaeology: a computerized simulation of reasoning by means of an expert system. *Computers and the Humanities* **19,** 37–52.

Langlotz, C.P. & E.H.Shortliffe 1983. Adapting a consultation system to critique user plans. *International Journal of Man-Machine Studies* **19,** 479–496.

Marshall, C.C. 1987. Exploring representation problems using hypertext. In *Proceedings of Hypertext 1987,* 253–268.

McCall, R., G.Fischer & A.Morch 1990. Supporting reflection-in-action in the janus design environment. In *The Electronic Design Studio,* McCullough, M. (ed.), 247–259, Cambridge, MA: Massachusetts Institute of Technology Press.

Miller, P.L. 1983. ATTENDING: critiquing a physician's management plan. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAM-5* **5,** 449–461.

O'Keefe, D.J. 1977. Two concepts of argument. *Journal of the American Forensic Association* **13 (3),** 121–128.

Patel, J. & A.Stutt 1988. Reducing uncertainty in archaeological interpretation using expectation based reasoning. In *Research and Development in Expert Systems V,* Kelly, B. & Rector, A. (eds). Cambridge: Cambridge University Press.

Patel, J. & A.Stutt 1989. Beyond classification: the use of artificial intelligence techniques for the interpretation of archaeological data. In *Computer applications and quantitative methods in archaeology 1989,* Rahtz, S. & Richards, J. (eds), 339–347. Oxford: British Archaeological Reports.

Rahtz, S., L.Carr & W.Hall 1989. New designs for archaeological reports. *Science and Archaeology* **31,** 20–34.

Reichman, R. 1985. *Getting computers to talk like you and me.* Cambridge, Mass.: Massachusetts Institute of Technology.

Rowland, R.C. 1987. On defining argument. *Philosophy and Rhetoric* **20(3),** 140–159.

Schiffer, M.B. 1976. *Behavioral Archaeology.* New York: Academic Press.

Searle, J.R. 1980. Minds, brains, and programs. *The Behavioral and Brain Sciences,* **3,** 417–457.

Shennan, S.J. 1989. The excavation of an early bronze age settlement at St. Veit-Klinglberg, Land Salzburg, Austria: an interim report. *Antiquaries Journal* LXIX, 205–224.

Shennan, S.J. 1990. Why should we believe archaeological interpretations? In *Interpretation in the Humanities: Perspectives from Artificial Intelligence,* Ennals, J.R. & J.-C.Gardin (eds), 80–100. London: British Library.

Smolensky, P., B.Bell, B.Fox, R.King, & C.Lewis 1987. Constraint-based hypertext for argumentation. In *Proceedings of Hypertext 1987,* 215–245.

Stefik, M., G.Foster, D.G.Bobrow, K.Kahn, S.Lanning, & L.Suchman 1987. Beyond the chalkboard: computer support for collaboration and problem solving in meetings. *Communications of the ACM* **30(1),** 32–47.

Stutt, A. 1988. Second generation expert systems, explanations, arguments and archaeology. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 353–367. Oxford: British Archaeological Reports.

Stutt, A. 1989. *Argument in the Humanities: a knowledge-based approach.* Milton Keynes: The Open University.

Stutt, A. 1990. Argument support programs: machines for generating interpretations. In *Humanities and the Computer: New Directions,* Miall, D. (ed.), 71–81. Oxford: Oxford University Press .

Stutt, A. & S.Shennan 1990. The nature of archaeological arguments. *Antiquity* **64,** 766–777.

Tilley, C. 1989. Excavation as theatre. *Antiquity* **63,** 275–280.

Toulmin, S. 1958. *The Uses of Argument.* Cambridge: Cambridge University Press.

Toulmin, S., R.Rieke & A.Janik 1979. *An Introduction to Reasoning.* New York: Macmillan.

Walton, D.N. 1989. *Informal Logic: a handbook for critical argumentation.* Cambridge: Cambridge University Press.

Winograd, T. 1988. Where the action is. *Byte* **December,** 256–258.

Winograd, T. & F.Flores 1986. *Understanding Computers and Cognition.* New Jersey: Ablex Publishing.

Woods, D.D. 1986. Cognitive technologies: the design of joint human-machine cognitive systems. *AI Magazine* **6(4),** 86–92.

Communication

19

From virtuality to actuality: the archaeological site simulation environment

BRIAN MOLYNEAUX

Introduction

The reality of archaeology as a practical activity and intellectual discipline is being extended and challenged by the development of graphics-oriented computer simulations. A simulation represents some aspect of the complex reality of an actual environment, generally for the purpose of problem-solving; a user commonly faces a set of problematic conditions and devises a strategy to overcome them. As such a situation can be endlessly recreated, the problem can be attacked in different ways over and over again until the user finds a satisfactory solution. The use of simulations has the potential, therefore, to improve practical and interpretive skills in the real world. Simulated environments have recently been developed for the teaching of excavation techniques, where they provide some significant advantages over traditional training excavations and field schools. Archaeological evidence is generally scarce and complex and there may be practical, social or ethical problems surrounding the disturbance of a site or the removal of artefacts; the excavation of a simulated site can potentially avoid such problems, providing a creative learning environment without the need for the violation or destruction of cultural property.

A good basic example of an archaeological excavation simulation is Walsh's (1981) ARCHSIM computer program, developed at the University of Minnesota. Although the graphics of the time were not interactive, the user was able to examine a site map on the screen and choose where to excavate by entering coordinates from the keyboard; and a cost was set for each excavation and analysis, so that with the limited budget provided, the user was encouraged to plan carefully.

A more advanced excavation simulation, one that takes advantage of modern interactive graphics, was developed in the Department of Archaeology at Southampton University (O'Flaherty 1988; Rahtz 1988; Wheatley 1989; Rahtz, Carr & Hall 1990; Wheatley 1991). The program, called *SyGraf*, is particularly concerned with the decision-making processes involved in excavation. Like ARCHSIM, it creates a simulated archaeological site accessible through the computer screen and requires the user to excavate it, choosing where to dig, what size and depth each trench should be and what type of excavation tool should be used, with the decision influenced by the size of an operating budget defined at the outset. With the implementation of interactive graphics routines, however, SyGraf significantly increases the potential for dynamic movement across the simulation environment and, hence, enhances the simulation effect.

Archaeological site simulations are also being developed using three-dimensional and solid modelling techniques (eg. Fletcher & Spicer 1992; Reilly & Shennan 1989; Reilly 1990). In these simulations, graphical objects can be examined from different points of view, creating the illusion that one can move through a site rather than having to view it from above as in plan-based simulations such as SyGraf.

The use of simulation in teaching has a price, however. Clearly, a simulated environment is not the same as the environment it simulates and, therefore, it is reasonable to be cautious about its appropriateness. This is especially so with regard to computer simulations, because their information is entirely metaphorical, unlike the sand, earth or other materials used in the first artificial sites constructed in laboratories (eg. Chilott & Deetz 1964). Since the computer simulation has no material relation to an actual archaeological site, its link to actual archaeological activity and the relevance of the experience it provides must be of primary concern.

It may be claimed that there is no substitute in archaeological training for direct experience at an actual archaeological site (cf. Wheatley 1991)—patently true, when taken in regard to the actual retrieval of artefacts and other information in excavation. As Reilly (1992) has suggested, however, graphical excavation simulations can provide new ways of seeing the archaeological record that bring into question some basic archaeological concepts, such as the definition of a site and the purpose of excavation itself—and this reassessment can be extended to the nature of archaeology itself, and its connection with the material world. Although simulations may not necessarily help the archaeology student to wield a trowel, they may be used to explore basic problems in archaeological thought that ultimately affect both excavation strategies and interpretations.

The goal of this chapter is to explore the essential issue of the link between excavation and its simulated counterpart, in order to provide the necessary theoretical background to support the implementation and further development of computer-based simulation environments in archaeology.

Reality and virtuality

A key point to establish in the assessment of simulations is that the seemingly obvious difference between 'real' and 'simulated' archaeological excavation is not as conclusive as it appears.

Interpreting the simulation as a fiction presumes that the reality of archaeology must be the material world encountered in excavation and the analysis of material objects. This view is almost forced upon us by the daunting complexity of archaeology, data recovery is affected by material and ideological conditions—the site environment, the research design, the tools and techniques used, and by the perceptions and actions of individual excavators—and yet, it seems that out of this loosely controlled process at least one has the material evidence recovered, and so, the foundation of the main archaeological enterprise. However, archaeology, like all human activity, is materially and culturally selective, the result of the fact that humans perceive and act within specific material and social situations. Consequently, the artefacts and other material objects recovered from an excavation do not have a singular identity—they are always viewed from an ideological position.

The socially relative nature of specifically *archaeological* information (as opposed to the physical attributes of material things) suggests that the reality of archaeology is not in the excavation site, regardless of the fact that it is the main source of raw data. Archaeology exists in the interface between the site and the excavator or analyst. It is the information itself, the metaphorical world created from the results of excavation: the inventories, databases and texts, the maps, diagrams and models and all the other information resources derived from the primary material (cf. Trigger 1989, p. 16).

The metaphorical basis of archaeological reality may be seen in the treatment of artefacts. Artefacts are generally only of use until they are classified and described, and then the descriptions take over in the analytical process. When their role as information sources ends, the objects are most often redeposited in new archaeological sites—the storage rooms of universities, museums or other institutions—except for a selection of 'representative' objects which become icons for the purposes of display or other intellectual commerce.

In contrast to reality claims for archaeology, it is common to regard the graphic display on a computer screen as essentially artificial, an image that has only a metaphorical relation to reality.

It is perhaps difficult to see the computer environment any other way. A computer user is usually deskbound, seated in a relatively fixed position as if in front of a window, and retrieves information that is, in the instance of site simulations, a representation of some other reality, conveyed with an intense visual bias and highly compressed in time and space. And yet, especially with regard to computer programs that provide the illusion of movement through physical space, as in hypertext (Martin 1990; McAleese 1989) and three-dimensional graphics environments (Wood & Chapman, (with contributions by Delooze & Trueman) 1992; Dew, Fernando, Holliman, Lawler, Malhi, & Parkin 1990; Lavender, Wallis, Bowyer & Davenport 1990; Reilly 1992), computer activity has an element of reality about it. This somewhat ambiguous or anomalous reality is described in computer theory in terms of *virtuality*.

Virtuality was first applied to computer environments by Nelson (1987) to describe computer processes and conditions that have some relation to external processes and conditions: 'Reality is that which is, *virtuality* is what seems to be' (Nelson 1987, pp. 2/7–2/8; cf. Reilly 1990, p. 84). The term *virtual reality* has been popularly adopted to describe the strong illusion of user participation in three-dimensional computer space. However, the difference between *real* and *virtual images,* a distinction that comes from the separation in physical optics of direct and secondary perception (see, for example, Gibson 1966, p. 227) is not actually relevant to the question of understanding the reality of the simulation environment. Although the debate concerning the reality of images is an important one in the critical analysis of interpretations in art (Gombrich 1960; Ucko 1977), the fact is that whether perceptions are defined as 'real' or 'virtual', they all consist of patterns of reflected light. This comparability suggests that the *physical* experience of perceiving and using such ostensibly different environments is similar. As Gibson observes, 'the same stimulus coming to the eye, however produced, will always afford the same experience' (Gibson 1966, p. 228). This democratization of perception, blurring some of the distinction between material objects in the world and representations of them, is significant here because it indicates that regardless of the limited range of information in a simulation environment, the perceptual experience of simulations can be of the same nature as perceptual experiences in the 'real' world. It is not the question of whether a stimulus comes from a first- or second-hand source, but how it *engages* the perceiver.

Indeed, when reality claims are set aside, it could be contended that the most significant difference in perceptual activity between external environments and those behind monitors—the scale of operation aside—is that the participant moves through the external one, whereas in the conceptual interior of the screen program, it is the environment that moves. In both, however, the user is physically and ideologically active, scanning the sensory environment for information and, with the aid of keyboard and mouse in the computer world, moving around and through it.

Computer site simulations may be far from an *earthy* reality, but their graphic displays may be considered as actual environments, explored by means of real-life activity and, therefore, suitable for the solving of archaeological problems. In this sense, the users are not external to the simulation, but part of it. Even if they cannot literally ‘step into’ the simulation screen, they are conceptual inhabitants, able to move metaphorically in virtual space and, with increasing facility in three-dimensional simulations, through it. Most importantly, the information recovered is much closer to the *real* archaeology of analysis and interpretation. A small, coloured shape in a graphic display on the computer screen, representing a sherd of pottery, is no more distant from the material world than a description of the find in a field report.

The learning environment

When archaeological site simulations are properly conceived as different forms of reality, instead of lesser imitations, their potential in learning may be better exploited. Because simulation activity is both physical and ideological, it is most important when simulations are designed and applied to take into account the entire interactive context of user and simulation. This approach may be described as a study of the *ecology* of the simulation environment (cf. Gibson 1966) and the potential for freedom of action within it. Using such criteria, some forms of computer-aided learning may be seen as deficient in the way they influence the actions of the user. For example, the value of computer-aided learning is often identified in the clarity in methods and goals that computer programs afford, an order simply enforced by the constructed nature of computer environments and capable of being masked by writing learning programs as games. Although games are engaging and may have some spontaneity in movement, the object is most often to achieve a specific end—to win (for games oriented more towards the learning process, see, for example, Castellan (1987) on teaching the scientific method).

Learning is not simply the attainment of knowledge or the achieving of a specified goal, however, but the *encounter* between the individual and the information environment, whether an archaeological site, a classroom, a book, or a computer monitor. As Rahtz, Carr & Hall suggest in regard to teaching in archaeology, students are not ‘consumers of received knowledge’ but ‘active participants in the learning process’ (Rahtz, Carr & Hall 1990, p. 193). Consequently, although simulations need to have qualities of play (see Wheatley 1991, p. 10), they also need to be as free as possible from the determined procedures and goal orientation of games.

To make an archaeological excavation simulation compatible with the dynamics of the learning environment, the user must be able to navigate within it without being entirely bound to specified pathways, as typified by menu-driven routines or the buttons and links of a hyperspace environment (see Kibby & Mayes 1989 for a discussion of how the use of explicit links to organize information in hyperspace biases navigation). User control is essential, as the first problem of an excavation is the fateful decision regarding where and how to dig. As the excavation strategy determines the nature of the information that follows, it should be part of the simulation routine. Among teaching simulations, SyGraf in particular has created an excavation environment which allows this basic freedom of movement. Through a more detailed discussion of SyGraf, it will be possible to consider the extent to which such computer-bound activities are adequate as a way of representing archaeological thought and practice.

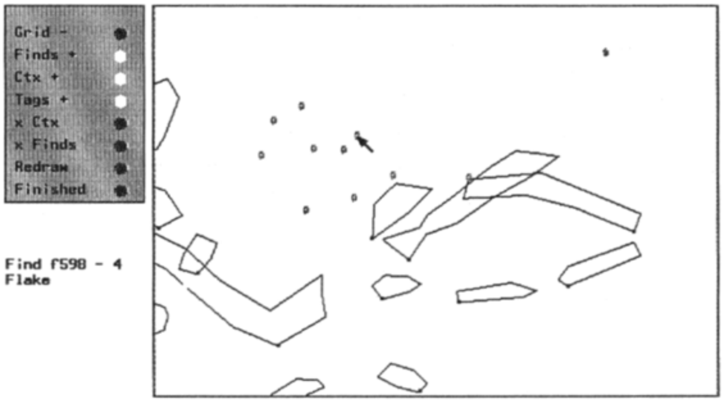
SyGraf

SyGraf divides the basic aspects of archaeological practice, information recovery and analysis, into two distinct information modes. On-site activity is simulated through the graphic display and the results of excavation are recorded in a series of relational databases.

The surface of a SyGraf site is a featureless green field that takes up most of the screen. Beneath the surface, the site consists of a series of two-dimensional plans, equivalent to the conventional floor plans of an actual excavation. Each plan represents a single occupation, and so a multi-component site may be conceived as a sequence of separate plans, placed one above the other, with the order of the levels intended to represent differences in depth and time and, hence, the separation of archaeological components. When a trench is excavated, any finds or features discovered within the area exposed are extracted from the master database and stored in a set of databases specific to the individual user. The finds and features are then displayed *in situ:* the finds, as icons representing their types, and the features, as closed polygons (Fig. 19.1).

The screen display is dynamic: the user can magnify areas of a site to obtain a close view of a single cluster of artefacts or draw back and view the site as a whole. Screen activity also extends to analysis. The user may identify artefacts while still working on the screen, as clicking the mouse pointer on each find and on a tag on each feature retrieves a description from the database. Most importantly, however, the graphic display is linked to a filtering routine in the database. The user may set a filter condition that will selectively display specific types of finds or features, according to description or excavation level (Fig. 19.2).

The dynamic nature of the SyGraf excavation environment provides an advantage over menu-based simulations. In the menu-driven archaeological site simulation *Fugawiland* (Price & Kolb 1986), for example, the excavator has no control over



Find f598 - 4
Flake

Figure 19.1 SyGraf excavation.

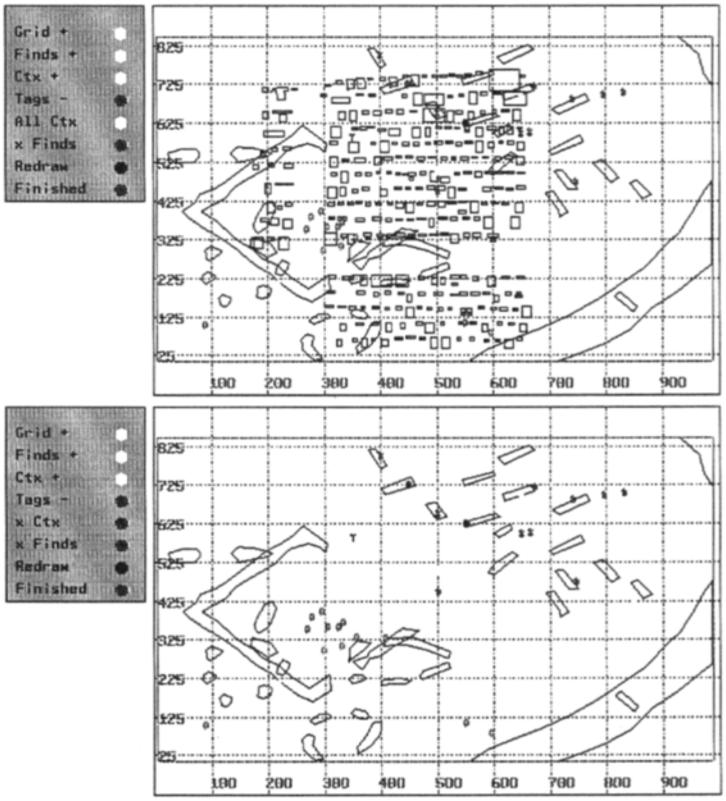


Figure 19.2 SyGraf excavation, showing full site, and site with filtering of data.

the location, size or extent of excavation at a site. These actions are initiated through menu selection and only the results of a total excavation are displayed. In the SyGraf environment, on the other hand, different kinds of sampling strategies may be implemented. The excavator may either specify the excavation coordinates on the keyboard or draw out a trench with the mouse, depending on the degree of metrical control desired. Within the SyGraf selection conditions, therefore, the individual is free to assess the situation and act accordingly: choosing an excavation area, activating the information retrieval process, or viewing and analysing the recovered information on the excavation screen and in the database. SyGraf has no necessary pathways— only necessary procedures. It is the difference between walking along a trail and in an open field.

At the same time, this freedom of action is constrained by the fact that excavation is subject to cost. The ostensible reason for introducing cost is to reinforce the fact that archaeology is part of an economic reality. At the same time, however, cost may also be regarded as a symbolic constraint that stands for all the constraints on excavation —time, materials, logistics, and personnel. With limited resources, the excavator must develop an excavation strategy in order to retrieve information from the site efficiently.

The way that SyGraf enables excavation and analysis is a complex mixture of conceptual and practical realities. It combines a metaphorical excavation routine with an analysis that is the equivalent of the methods used at an actual site. When the archaeological site is displayed, the user is in the position of a remote observer, who is capable of seeing the entire excavation in a single glance, but is able to carry out work on its surface. When the information has been recovered, however, the user becomes a participant, the equivalent of the field archaeologist, sitting in the field tent, office or laboratory, assessing and sorting the results of the excavation.

These features provide an interactive and open-ended task setting that is particularly useful in the training of archaeology students. Within the constraints of the simulation, each action and sequence is unique to the individual user; as the simulated excavation may be repeated, however, individuals may be assessed on the same range of material. In addition, the use of the scaled-down environment of a computer means that immediate feedback is possible; the substantial reduction of the gap between planning, information recovery and analysis therefore enhances the potential for experimentation and practice in analytical techniques. And on the whole, these procedures will make students more aware of the effect their decisions may have on the recovery and interpretation of archaeological data (cf. Rahtz 1988; Wheatley 1991).

The evaluation of results

The use of simulations creates new responsibilities in the teaching of archaeology. Most importantly, it exposes the problem of access to information and what effect this has on the analysis and interpretation of a site—and, hence, on the assessment of students' results.

The student works with the limited information available from partial excavation, but the instructor has access to the total information from the site. If students are assessed by a comparison of their findings with an *authoritative* interpretation, however, then the simulation becomes an implicit game. Students do not compete directly— but winners and losers are created by the extent to which their interpretations agree with this interpretive ideal. This form of assessment is contrary to the philosophy of SyGraf, as it supports the illusion that excavations yield concrete facts and relations which the student must learn to recognize. In so doing, it ignores the fact that information is relative to the perception of individual users and to the specific circumstances of their actions—and, therefore, that there is no single correct answer or interpretation. Indeed, with the freedom of choice in the SyGraf environment, one would expect a wide variation in information retrieval and interpretation from equally competent excavation routines, simply because some excavators found more artefacts and features than others. It is therefore the course of decision-making, rather than the inventory of artefacts, that distinguishes the abilities of the excavators in the simulation. This dynamic interaction is lost if only the results of excavation are stressed.

When assessment becomes more focused on the individual than the answer, each activity in the simulation environment becomes part of an historical process—a simulated history. By examining the results within such specific circumstances, the instructor is better able to judge how the student has responded to the contingencies of the situation. Each excavation should therefore be evaluated in its own terms, according to how the student develops an excavation strategy and pursues it.

This focus on the individual does not, however, destroy the possibility of consensus. For when the individual excavations and interpretations are brought together, they constitute an archaeological discourse about the site. Such a discourse, with its multitude of approaches and ideas, will be closer to a social reality than the interpretation represented by any single view—including the privileged one that exists for the site as a whole—and, therefore, it has the potential to represent a real context for learning.

The locus of reality in simulations

The combination of observation and participation gives the simulation its potential as a learning device, in that it attempts to create a setting in which decisions consistent with actual archaeological practice are made. On the basis of these criteria, however, it may appear that a two-dimensional map-like simulation such as SyGraf is less useful in conveying the dynamics of archaeological practice than the more sophisticated three-dimensional models, such as Reilly's *Grafland* (Reilly 1992). The Grafland environment may be excavated in depth as well as surface, with the result that finds and features may be represented within a matrix of relations, or contexts, in space and time. Such a model provides a significant potential for the exploration and analysis of site formation processes and proxemic (socio-spatial) relations.

However, archaeological sites are never actually seen as the volume of information pictured in the three dimensions of a graphical illusion: a site is always a sequence of opaque surfaces, no matter how represented (cf. Reilly 1990, pp. 84, 87; Wheatley 1991, p. 10). From this perspective, the value of SyGraf is that it is able to represent important aspects of the actual *situation* at an excavation; it attempts to simulate problems that the archaeologist faces *during* excavation, rather than the problems faced when the excavation is completed and the results are being analysed.

The unique power of the three-dimensional simulation, on the other hand, is precisely in its departure from reality, its ability to model site volumes and disassemble them selectively to expose relations in space and time that were never visible in reality. Three-dimensional simulations are therefore closer to the metaphorical world of archaeological interpretation, in that they construct the kind of imaginative pseudoreality necessary to explore the complexity of site occupations and site formation processes over time.

Because it is two-dimensional, SyGraf is incapable of such manipulation; its surface represents the ground of any archaeological environment. It has an immediacy that serves a different kind of heuristic purpose, rather than a diminished one. It is the archaeologist working with a site map, rather than the omniscient observer and participant of Grafland.

Conclusion

Whether constructed in two or three dimensions, simulations challenge archaeology in that they bring into question the nature of archaeological information and its relation to material and social realities. Most clearly, they show that although the computer in an educational context is regarded as a teaching tool, it is a dynamic environment, and teachers must therefore understand the ecological relations involved in its use to be effective.

In regard to SyGraf, the exercise is not simply a routine for the extraction of information. SyGraf is consistent with actual archaeological practice in that the selectivity in the excavation routine ensures that analysis may lead to reasoned hypotheses, but will never reach perfect conclusions. This is in contrast to teaching programs that are, in essence, animated books, and games, in which users are free to select information pathways but have no control over them. Such programs offer answers, but miss out on the conceptual strength of the simulation environment.

An ecologically aware simulation such as SyGraf may be easily integrated into an existing archaeological curriculum, where it may be used as an interactive resource for the general analysis and interpretation of excavations, or for more specific studies, such as the comparison of sampling strategies or the analysis of the relation between artefact distribution and density and other site features.

What is most important about such simulations is that they will take on the character of the research and teaching programme within which they are used, rather than imposing the structure, point of view and information of the program's source. They provide the exciting potential to create a dynamic interactive environment, one in which instructors are challenged to develop and implement teaching strategies as much as their students are challenged to explore and analyse archaeological landscapes.

References

Castellan, N.J. 1987. The use of computer-based games in teaching research methodology. *Collegiate Microcomputer* **5,** 209–12.

Chilott, J.H. & J.Deetz 1964. The construction and use of a laboratory archaeological site. *American Antiquity* **29,** 328–337.

Dew, P.M., L.T.P.Fernando, N.S.Holliman, M.Lawler, R.Malhi, & D.Parkin 1990. Illuminating chapters in history: computer aided visualisation for archaeological reconstruction. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 15–25, & 117 (plates). Winchester: IBM UK Scientific Centre.

Fletcher, M. & D.Spicer 1992. The display and analysis of ridge-and-furrow from topographically surveyed data. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 10. London: Routledge.

Gibson, J.J. 1966. *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Gombrich, E.H. 1960. *Art and illusion: a study in the psychology of pictorial representation.* London: Phaidon.

Kibby, M.R. & J.T.Mayes 1989. Towards intelligent hypertext. In *Hypertext: theory into practice,* McAleese, R. (ed.), 164–172. London: Intellect Limited.

Lavender, D., A.Wallis, A.Bowyer & P.Davenport 1990. Solid modelling of roman Bath. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 7–13 & 116 (plates). Winchester: IBM UK Scientific Centre.

Martin, J. 1990. *Hyperdocuments and How to Create Them.* New Jersey: Prentice Hall.

McAleese, R. (ed.) 1989. *Hypertext: theory into practice.* London: Intellect Limited.

Nelson, T.H. 1987. *Literary Machines.* Swathmore: T.H.Nelson.

O'Flaherty, B. 1988. The Southampton-York archaeological simulation System. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 491–498. Oxford: British Archaeological Reports.

Price, T.D. & M.J.Kolb 1986. Fugawiland. Test pilot 1.00.

Rahtz, S.P.Q. 1988. A resource based archaeological simulation. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 473–90. Oxford: British Archaeological Reports.

Rahtz, S.P.Q., L.Carr & W.Hall 1990. Creating Multimedia Documents: hypertext-processing. In *Hypertext: state of the Art,* Green, C. & R.McAleese (eds), 183–192. Oxford: Intellect.

Reilly, P. 1990. Towards a virtual archaeology. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 83–92 & 119–120 (plates). Winchester: IBM UK Scientific Centre.

Reilly, P. 1992. Three-dimensional modelling and primary archaeological data. In *Archaeology and the Information Age*, Reilly, P. & S.Rahtz (eds), chapter 12. London: Routledge.

Reilly, P. & S.Rahtz 1990. *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990.* Winchester: IBM UK Scientific Centre.

Reilly, P. & S.Shennan 1989. Applying solid modelling and animated 3-dimensional graphics to archaeological problems. In *Computer Applications and Quantitative Methods in Archaeology 1989*, Rahtz, S. & J.Richards (eds), 157–66, Oxford: British Archaeological Reports.

Trigger, B.G. 1989. *A History of Archaeological Thought.* Cambridge: Cambridge University Press.

Ucko, P.J. (ed.) 1977. *Form in Indigenous Art.* London: Duckworth.

Walsh, V.A. 1981. Computer-aided instruction in archaeology. In *Computing in the Humanities*, Patton, P.C. & R.A.Holoien (eds), 215–229. Aldershot: Gower.

Wheatley, D.W. 1989. *SyGraf—A Prototype Graphically Based Excavation Simulation for Computer Aided Learning.* Southampton: Unpublished MSc thesis, University of Southampton.

Wheatley, D.W. 1991. SyGraf—resource based teaching with graphics. In *Computer Applications and Quantitative Methods in Archaeology 1990*, Lockyear, K. & S.Rahtz (eds), 9–13. Oxford: British Archaeological Reports.

Wood, J. & G.Chapman (with contributions by K.Delooze & M.Trueman) 1992. Three-dimensional computer visualization of historic buildings—with particular reference to reconstruction modelling. In *Archaeology and the Information Age*, Reilly, P. & S.Rahtz (eds), chapter 11. London: Routledge.

20

The electronic capture and dissemination of the cultural practice of Tibetan Thangka painting

RANJIT MAKKUNI

Introduction

To glance for the first time at the painting illustrated in Figure 20.1, is to be transported to an evocative world charged with fantasy and symbolism. The image depicted is not a portrait, but a symbol of Sakyamuni, Buddha, or the Enlightened One. The painting is called Thangka (Tucci 1949; Pal 1983). Typically, images on Thangka paintings, like those expressed in stone and metal, represent the various deities that populate the Tibetan artistic pantheon. Blazing with symbolism, their purpose is to transport the viewer into the supernatural and paradisiacal world of the deity, a world conceptualized by theologians and visualized by artists for countless generations conveying the Tibetan Buddhist ideal of self-realization and harmony with the cosmos. The paintings are invested with special powers, and are dedicated in an attempt to secure both material benefits and spiritual well being for the donor, the donor’s family, and the greater community (Pal 1984b).

Visually expressive and stimulating as these images are, the depiction of the various deities on Thangkas is regulated by very precise rules of composition. Some of these rules of composition were represented in canonical treatises describing verbally the basic iconography of the various deities. Other rules were expressed visually through proportioning diagrams and example sketches of deities (see Fig. 20.2). The sketches demonstrate, for example, various compositional schemes of Thangkas, the postures of deities, the gestures that deities make with their hands, the symbolic objects that their hands clasp, the rendering of the deities’ garments and accessories, and landscape elements that echo the deities’ spiritual qualities. These rules of composition are timeless as they were passed down from master to pupil through successive generations, the medium of transmission consisting of proportioning diagrams, example sketches, and verse in Tibetan canonical treatises.

A variety of forces threaten the practice of Thangka painting, moving the craft towards irrevocable simplification. The Electronic Sketch Book of Tibetan Thangka Painting project (Makkuni 1989b) is conceived as a way of using interactive computing and video technologies to: preserve in electronic form Thangka expertise and the cultural context in which Thangkas are created; use the records in educating people about the world of Thangka painting. The project is a creative collaborative effort between groups: systems designers of Xerox PARC, art historians and Tibetologists of the Asian Art Museum of San Francisco, a renowned Tibetan monk, and a master Thangka painter.

The electronic sketch book is seen as having two roles: a preservation role, and a dissemination role. In the preservation role, the sketch book takes form as a chronicle, an audio-visual diary of Thangka imagery and expertise similar to traditional manuscript illuminations and narrative paintings. In the dissemination role, the sketch book serves as a medium of transmission, one that can connect Thangka masters with beginning students, and, in museum settings, serve as an interpretive guide to both the cultural context and process of Thangka paintings. On a philosophical note, the sketch book relates the sense of ‘contemporary time’ with the ‘traditional time’ of Tibet.

The introduction of a new communication and representation medium is bound to affect the practice of Thangka painting. In this chapter, we present our motivations for the use of electronic technologies for preserving and disseminating Thangka painting. We examine the relationship of the art to the rich cultural practice, illustrating the relationship between media, ritual and the designed artefact, and point out the limitations and benefits of electronic technologies in the capture and transmission of Thangka painting. Focussing on process as well as product, our research illustrates themes of characterizing the design expertise behind the artefact, and the electronic transmission of cultural presence. Specific to how people interact with electronic media, we address expressive means to interact with computers, and the re-creation of the cultural experience using electronic media.

The project broadly envisions the supportive role of electronic technologies to the preservation of deteriorating traditions. In applying electronic technologies to Thangka painting, the customers of the research are members of the Tibetan community, scholars interested in preserving and disseminating Asian and Tibetan art, researchers and designers of electronic systems, and the general public at large. The space of interested parties for the project is large, and the appearance of the sketch book will differ according to the audience. We envision scenarios for incorporating electronic technologies into the Tibetan cultural setting. Keeping in mind the broader vision for the use of electronic technologies in the world of Thangka

Figure 20.1 Thangka painting of Sakyamuni Buddha, sixth century BC founder of Buddhist teachings.

painting, we describe a particular implementation of the sketch book for a western museum educational setting—the galleries of the Asian Art Museum of San Francisco. For the museum setting, the focus of the sketch book is on the preservation of the cultural practice of Thangka painting, and public dissemination aimed at the general museum-goers who have no prior experience with Tibetan art, and who have little or cursory knowledge of computers.

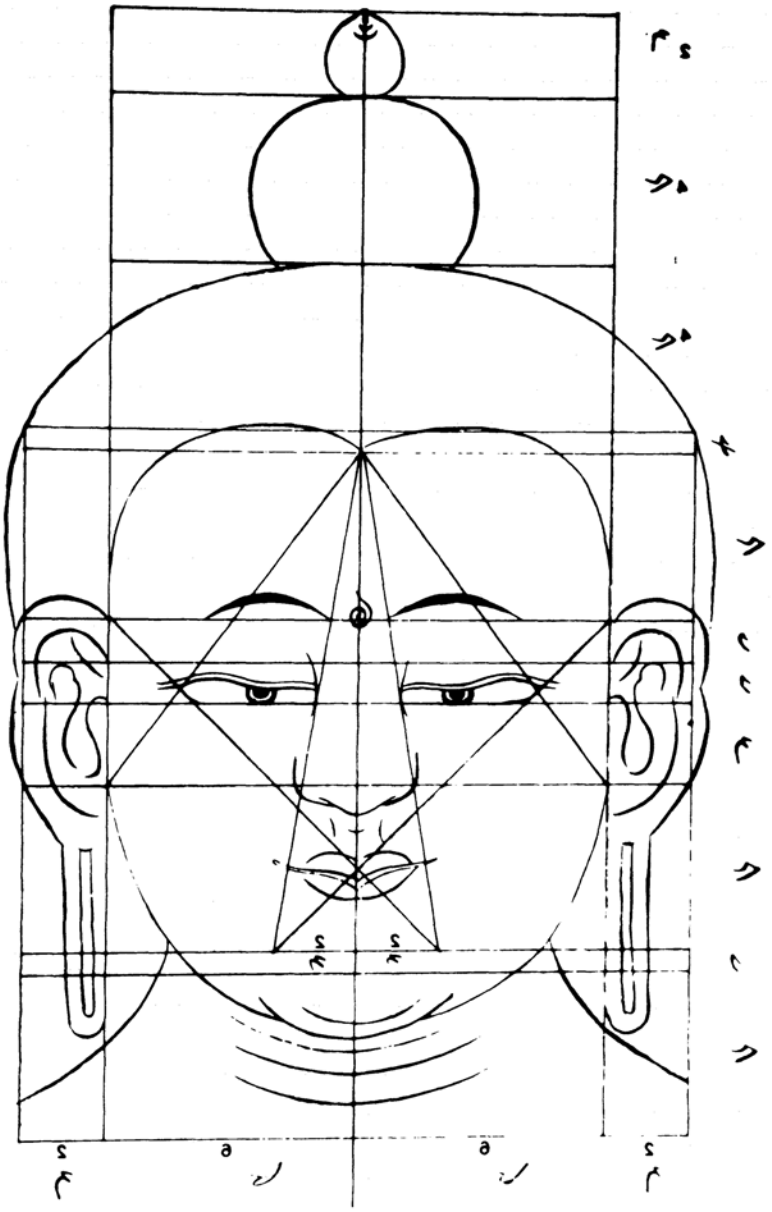


Figure 20.2 Proportioning diagram of Buddha's face.

The sketch book takes form as a museum installation by which general museum-goers can gain a glimpse of the world of Thangka painting—its sounds, sights, processes, and cultural context. Associated with the general museum audience we faced particular challenges which shaped the form and content of the sketch book: for the general museum audience, the sketch book must be easy to use; its user interface model must be simple to understand; it must allow museum-goers open-ended explorations. With respect to the sketch book's construction, the presentation materials had to be authored by museum curators who had little exposure to computers.

The sketch book contains a video database of sounds and images of the process of Thangka composition including: images from the museum's rare painting collection, images of rules of composition, techniques, and Tibetan cultural life, live recordings of a master Thangka painter's composition process, curatorial analyses of paintings, and a monk's discussion of theology. Mediated by a computer system, this database is used in two modes: authoring and presentation. In the authoring mode, museum curators collect video source material of Thangka imagery, edit and condense the source material into video records, create a computational model of the video records in the database, organize the records, and create presentations of Thangka related topics for the public. In the presentation mode, an interactive graphic interface allows museum-goers to explore the curator's presentations.

The museum setting of the sketch book consists of a wall containing three video monitors and a computer screen (see Fig. 20.3). By pointing at and touching elements of a Thangka painting on the computer screen to play back video records, the

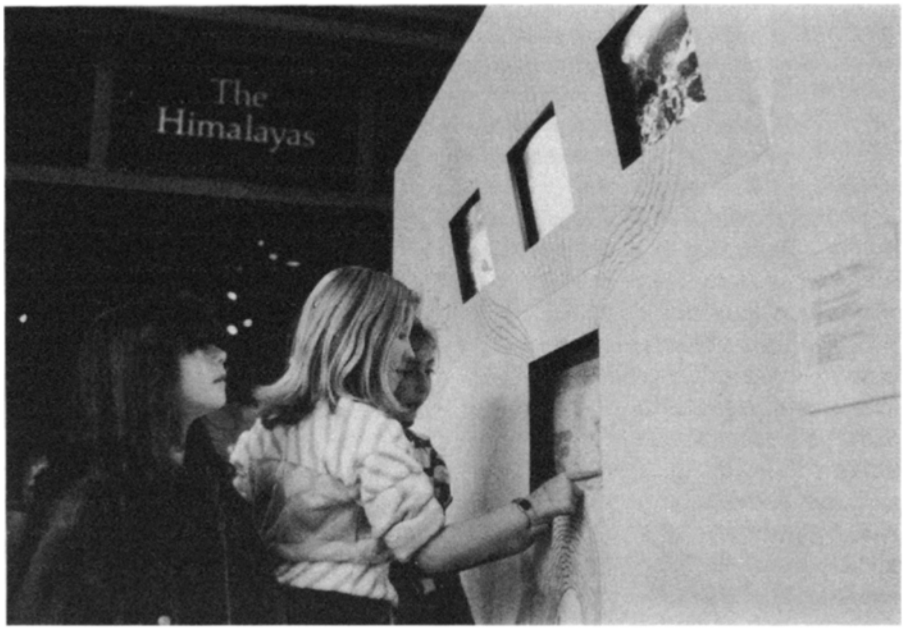


Figure 20.3 The Electronic Sketch Book in a museum setting.

museum-goer gains access to the different elements of the video database—paintings, live sketches, discussions, curatorial analyses.

Motivations

The word ‘Thangka’ literally means ‘Something that is rolled up,’ hence, a rolled-up image or a painting scroll. The Tibetan words for ‘to paint’ literally translate as ‘to write Gods.’ A cross between symbol, ornament, and design, the Thangka is a way of ‘seeing’ the world from both aesthetic and spiritual points of view.

The centre stage of a Thangka painting is usually occupied by a deity. In the painting shown in Figure 20.1, the Buddha, dressed in a monk’s robes, is shown meditating in still repose, and turning away from all attachments to seek a cure for the world’s sorrow. A lotus supports his body for, indeed, his feet cannot rest on the physical earth. His right hand extends over the length of his body and touches the cushion upon which he meditates, calling the earth to witness the occasion of his enlightenment. This expression is called the gesture of ‘earth witnessing’ or ‘earth pressing.’ Similarly, his left hand, expressing the gesture of ‘meditative equipoise,’ accepts in its palms a begging bowl. Positioned around the perimeter of the throne, although greatly reduced in scale, are acolytes and attendants, goddesses and demons, leaves and flowers, dragons and deer, whose purpose is to emphasize the majesty of the central personage. Various offerings—the symbolic wheel, conch shell, bowls, jewels, auspicious birds—are laid before the Buddha. Motifs of landscape, evoking mystery—steep mountains and canyons, turbulent lakes with dancing waves, and whirling clouds—form the background of the painting. These symbols of expressive gestures, auspicious offerings, and mysterious landscapes form a rich visual language, which is the medium of communication between the painter and the viewer.

Despite the bewildering complexity of the painting, the technical basis for Thangka paintings is a series of rectilinear diagrams (see Figure 20.2). The depiction of the various deities on Thangkas is ordered by very precise rules of composition, among them the theories of the bodily proportions of the various deities that make up the Tibetan artistic pantheon. These theories of proportion, handed down, generation through generation, from master to pupil, have been the means of transmitting the craft for the last two thousand five hundred years. These theories are not the work of one artist or generation, but the work of generations of craftsmen, the fruit of communal thought. According to art historian Coomaraswamy (1964):

> This communal thought is not only the popular thought, but that of the greatest and wisest minds seeking to impress their vision on successive generations. However there is a fatal weakness of communal art: it has no power to resist the corruption from without. It is beautiful by habit, rather than by intention, so a single generation under changed conditions is sufficient to destroy it.

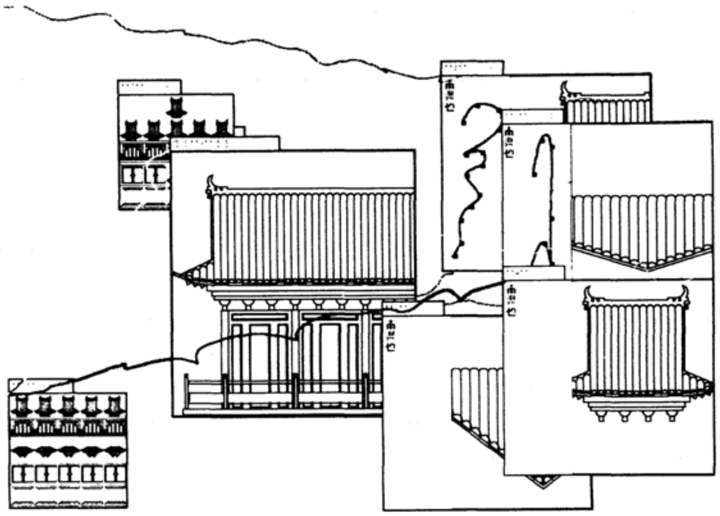


Figure 20.4 Temple designer exercising a gestural language to interactively compose and record the process of composing temple facades. Various threads tie up the different design activities over time.

The annexation of Tibet in 1959 by China led to the exodus of Tibetans from their homeland, among them their spiritual and temporal leader, the Dalai Lama. Tibetan monasteries, once the rich repositories of Thangka paintings and other cultural artefacts, were impoverished and many of their paintings destroyed. Many of the painters and other craftsmen were scattered in various refugee settlements all over the world. In a battle to preserve and reconstruct the cultural continuity, the Dalai Lama has asked painters and craftsmen to disseminate the traditions in the West. Thus the painters have become wanderers, traveling the world as a living archive. It is in these painters, more than in collections of Thangka paintings in museums, that we find the knowledge of the ‘ways of painting.’

The Electronic Sketch Book project suggests the potential of electronic technologies —computers and video—to preserve and disseminate long standing craft-traditions, such as Thangka painting, endangered by external forces. Museum exhibitions preserve the artefact of craft. Scholarly work in academic institutions illustrates skills and techniques and shows relationships between culture and craft. While there has been a considerable and exemplary effort in preserving artefact and illustrating techniques, our work offers some support for the community that practises the craft, assisting the human carriers of the ways of creating artefact. While craft has been disseminated by academic (literary) means, we recognize the need to incorporate—rather than exclude—the genius of the community of craftsmen in building tools for preserving and disseminating their craft. The best designers and judges of any tools that preserve and disseminate craft are the members of the community that actually practises the craft.

Our experience with the use of computing and video to support craftsmen and designers arises from projects that we have successfully completed (Makkuni 1987), in particular the composition of Chinese temple facades (see Fig. 20.4). We explored the possibility of representing successful design processes—directly representing the ‘process of formation’ of the artefact, not just the end artefact, in computational form—and of using the representations to preserve the successful processes. In other work in our laboratory (Harrison 1986; Stults 1986a; Stults 1986b), we have explored the representation of design processes, especially the communications of designers, as records on video tape and videodisc (see Fig. 20.5). In these studies we come to understand and demonstrate how electronic technologies can serve design groups, permitting the groups to keep their practice relevant in technologically and economically developed settings and, at the same time, permitting the groups to maintain their traditional human skills and relations.

Based on these explorations and studies, we believe that representations of process, when repeated and re-enacted across situations, connect members in a design group across time, communicate experience between members and across projects, and provide a basis for formalized design craft. In addition, the representations of design experience are valuable in educating people about the craft. They can provide beginning designers with a rich library of previously preserved scenes, which, in turn may be assimilated into future design practice. When viewed over long periods of time, design craft can be viewed in a state of flux: some in the process of formation, experimentation, simplification, or, as in the case of Thangka painting, in the process of deterioration.

Our support of Thangka painters in preserving and disseminating their craft is an extension of work in representing process. While traditionally Thangka painting has been transmitted by structured rules of composition, these rules are highly evolved and stable, they do not exclude artistic exploration.

The images of deities commonly depicted on Thangkas, along with their counterparts expressed in a variety of media—such as stone, metal, or wood sculptures, ink manuscript illustrations painted on paper and cloth, acrylic and water colour

Figure 20.5 Designers communicating and collaborating in video space.

paintings on cotton and silk banners, and mural paintings and inscriptions on the walls of monasteries—certainly illustrate the technical prowess of the artists in adhering to the rules of composition, but at the same time, also illustrate the artists' imaginative power of visualization. Though all artists obey the rules of composition, no two paintings or sculptures are alike. The many different executions of the same deity illustrate the delicate relationship between, first, the artists' technical virtuosity in the medium to express the theologians' visions, and second, the imaginative power of the artists in rendering these visions to reflect the artists' love, devotion and admiration of the deity being depicted. This relationship, between the remarkable conformity to basic iconography and the ethnic stylization of the deities, is also felt in the various countries where Buddhism flourished (Snellgrove 1978; Pal 1984a). The artist reproduces a design handed down unchanged over centuries in a practice of aesthetic fidelity to angelic prototypes. But, within the language of the craft, there is incredible freedom to improvise. These images, regardless of medium, time, and place, illustrate the artists' faithfulness to the well-established rules of composition, and ability to accommodate artistic exploration within those rules. This, notwithstanding the preservation and dissemination of the craft as it has been practiced in traditional media, makes even the expression of Thangkas by applying structured computing and video machinery a domain worthy of examination.

Admittedly, it would be scholarly conceit, blinded by optimism, to believe that, in bringing electronic technologies to the craft of Thangka painting, change is escapable. Electronic technologies herald change with the potential of both improvement and degradation of the process of craft. Amidst this dilemma of change, we cannot remain satisfied by avoiding the use of electronic technologies to preserve and disseminate the craft of Thangka painting. In any event, our use or avoidance is measured against the forces endangering the craft: rapid collecting of Thangkas, commercializing Thangkas, simplifying the process of creating Thangkas to meet the tourists' demand for mementoes, and painters abandoning their craft in search of economic opportunity. These forces are already moving the practice of Thangka painting toward irrevocable simplification, transformation, or degradation. Timeliness is important; delay might leave only an extinct craft. Hence, we propose to bring electronic technologies to Thangka painting now, and to do so with great reverence for the craft, as did countless generations of painters who produced these ageless, admirable paintings.

Cultural setting of Thangka painting

Craft is a communal process of making artefacts in which the designers are involved with the whole process of design, especially in the development of tools and supportive traditions, and acquiring of hand and mind skills (Smith & Lucie-Smith 1986, pp. 11– 40; Yanagi 1972, pp. 197–224). Historically, it has been identified with producing artefacts that were necessary for life and hence its beauty is born in communal use. The craft world of Thangka painting encompasses mythologies, symbolism, tools, skills, ritual enactment, physical setting, and artefact. The community includes the producers, the donors and the painters, and the consumers, the priests, devotees and ordinary Tibetan people.

The act of painting may be viewed as a ritual, through which, for a particular time and place on earth, the macrocosm is recreated through the microcosm of the performance of painting. ‘He who would draw a figure, cannot do so, if he cannot be it,’ (Dante), ie. the artist identifies with the theme to be painted. The artist becomes, for the duration of the ritual, the emanation of the deity that is to be painted. The Thangka master verbally invokes the deity, identifies himself with the deity, offering chants as he produces the drawings of paintings. The offering of chants, and the offering of painting strokes made steady and precise through years of exacting discipline are indispensable parts of the ritual. Thus the act is extremely rich, illustrating a mix of technique, chanting, song, breath, and setting: traditional boundaries between myth, painting, and song blend into one another, as they support each other. Thus, as a goal, the capture of the performance of painting must re‘present’ the theater, the place in which it is performed, and the enactment of the ritual through diverse communications—song, breath, prayer, painting techniques.

In its use, the Thangka represents a cosmic vision of the universe.The Thangka image is used as a support to meditation, and in ritual settings as props that transform ordinary space into a special theatrical place. The Thangka image in the ritual setting transports the viewers into the paradisiacal primordial time of the deity. Various sense offerings such as yoghurt (connoting taste), sandal paste (smell), and silk (touch), are offered by the Tibetan worshippers to symbolize the death of the ego. Thus in the offering, the Thangka is activated, and for the duration of the ritual, the worshipper becomes the deity. Even today, the presentation of the epic Mahabharatha on Indian television is a powerful example showing how intimate the connection between people and the image is, regardless of medium, time, and place. The ritual of offerings continue: Indian audiences garland their television monitors with flowers, mark the monitors with sandal paste. The central idea is that human beings have direct access through the ritual into the world of the divinities. The divine stories interact with day to day life of the people. Old rituals survive and are given new expression with the introduction of each new medium. The activation of the image in the human beings reflect a world view which ‘asserts inter-relationships, interconnections, ecological balances and thus the inter-dependence of the animate and the inanimate, and the possibility of the transmutation of matter into energy and energy into matter’ (Vatsyayan 1986, pp. 567–575).

The origins of the ritual of circumambulating the stupa and other sacred images reenacts the myth of tracing the different directions, the celestial, the ethereal, and the terrestrial. Participation in the ritual in the special place and time creates a theater in which the worshipper for the specific time and place becomes deified into icons. In building presentations of Thangka, especially in a museum setting, we become sensitive to the theater and performance, and understand how the ingredients of the performance can be retained, absorbed or transformed in contemporary settings. Transplanting rituals out of their original context raises controversy, since, if the ritual is brought out of context, then ritual efficacy might become transformed into entertainment. However, in the choice between proposing radical change and retention of old rituals, project members favoured a smooth transition between the old and the new.

Other traditions have also played important roles in transmission, such as, the theatrical tradition in which deities and historical figures portrayed on Thangkas are brought to life, the folkloric tradition in which the history of the figures are elaborated, and the meditative traditions in which Thangkas were put into use (*pers. comm.* Kohn).

As shown in Figure 20.6, the Thangka image relates to tools, skills, mythologies, rituals, and the physical setting. The diagram combines both the point of view of production and the use of Thangkas. From the viewpoint of production, the relationship between tools and the Thangka image includes, for example, knowledge of the materials, preparation of colours, canvas and brushes. Hand skills include composition skills, the skills in generating proportioning diagrams, preparing overall layout, and realizing an artistic idea in an image through technique. From the use viewpoint, donors of Thangkas appreciate the artist’s execution of the painting, and the mastery of technique. The relationship between mythology and the Thangka image includes the artist’s imaginative ability to translate an idea into image, to assert the artist’s individuality while remaining faithful to established rules of composition, and the viewer’s ability to be transported from the Thangka image into the world of the deities. In the construction of a Thangka, the artist through ritual evokes the presence of the deity. In use, the Thangka serves as a reference to the meditative rituals of the spiritual practitioner: by a process of comparison, the practitioner unifies the image in the mind with the external Thangka, and ultimately identifies with the spiritual qualities of the deities depicted on the Thangka. Finally, the Thangka image is intimately related to the physical setting in which the performance of painting and use occurs.

Electronic capture and dissemination of Thangka painting

The purpose of the electronic sketch book is to aid in the capture and dissemination of the cultural processes of Thangkas. The craft environment will still include actual paintings, but the purpose here is not to study or display the finished paintings, for museums and books do that well enough. Rather, the purpose is to introduce into the craft environment a medium that is fundamentally about process, and hence about the sense of time. A craftsman takes actions towards a Thangka, such as generating an element, identifying with the theme to be depicted, or composing a whole painting. Unlike the scholar or museum curator

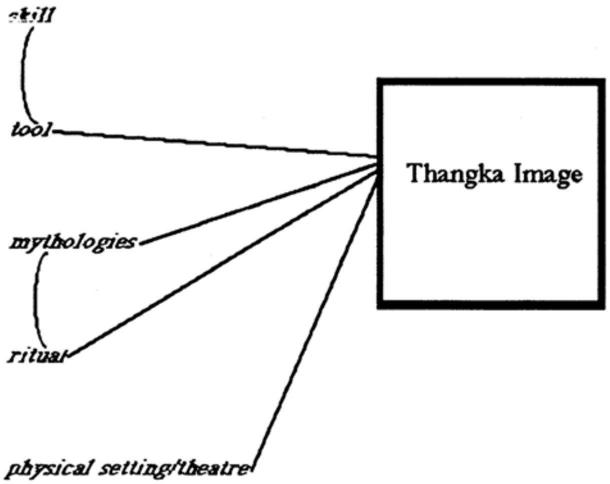


Figure 20.6 Relationship of the Thangka image to mythology, ritual, tool, and physical setting.

who examines and collects these finished paintings, we are concerned with representing and collecting the scenes of actions in which the craftsmen create the Thangkas, and tying those to scenes in the communal use of the Thangkas.

Through these scenes, each action, however large or small, whether it is the shading of a deity's eyes, the invocation of a deity, or examining the different compositional schemes for a whole painting, will be capable of re-enactment. By re-enacting process, aspiring painters or spectators can retrace and re-experience the actions of the Thangka master, and learn the craft through action. By replaying scenes showing communal use and discussions with Thangka practitioners and craft-persons, beginning students and museum-goers can gain an understanding of Thangkas in their cultural context, and learn how to interpret the Thangkas. The research challenge is to examine the nature of a craft process that is being renewed—and altered—using the electronic sketch book.

Thus the following questions can be put:

1. How can the visual world of the traditional media be brought into harmony with the visual world of colour monitors, and colour printers?
2. How can the craft-person acquire, experience, and subsequently impart the learning of hand skills using electronic technologies?
3. How will the electronic environment support the craft-person's evolving contribution to the craft?
4. Will electronic media become essential to the sustenance of the renewed craft?
5. Will the aura of electronic media influence the sensibilities of Thangka masters?
6. How will the master use the limitations or benefits of electronic media— precision, standardization, and replication, and homogenization—towards Thangkas?
7. Will the master absorb and express this aura in the paintings?
8. Can electronic media capture the rich cultural presence of Thangka painting?
9. Outside the traditional setting for Thangka painting, what is the experience of participation for a non-initiate, such as a western museum-goer?

Scenarios of using electronic media in the cultural setting

We now illustrate, through the use of sketches, scenarios in which electronic media, ritual practices, and physical setting interact in the overall cultural experience of Thangka painting, paying attention to how physical space appears with respect to the technology. This differs from the customary viewing of computers and video technologies as self-contained boxes. We consider both the worlds inside and outside of the box, the world of technological systems embedded in the social systems of use. In the following sketches, we present diverse situations for the Thangka experience, ranging from the private contemplative setting for a student engaged in painting, to the semi-private initiation ceremonies, to public performance settings in which deities and historical figures portrayed on Thangkas are brought to life.

In a private contemplative setting, a Thangka student painter uses computer and video technologies to connect with a master, and with records of Thangka practice. The records contain images, diagrams of rules of composition, folklore of a master elaborating historical figures and deities. By singing, chanting, and painting along with the master, the student connects to the master, and in turn, to the lineage of countless previous masters. The student controls the video playback of records using the computer and video for viewing the imagery. Since the aim of the education is to prepare the student to build self-confidence in the execution of strokes, we might critically ask: Does the freedom of re-playing video scenes of painting erode the student's confidence? The act of painting, like life, occurs once and for all, without deliberation, and afterwards no correction is permissible. The constant feedback of the master is required in the educational process, and hence the electronic systems may work best if they are used as references in support of classroom education.

In the theatrical setting, dancers interact with recorded imagery as part of the emerging choreography. Images of deities and historical figures displayed on the monitors are activated and brought alive by the dancers. The script of the dance allows the dancer to carefully interact with the recorded image. Through sensors on the podium, the dancer's footsteps trigger retrieval commands, which, in turn, summon and display recorded imagery on the video monitors. In terms of pure form, the dancer's posture can interact with previously recorded postures on the video screen, and can illustrate subtle variations of particular postures. If the performance itself is recorded and fed back onto the monitors, the dancers can layer improvizations on the previously recorded steps, and assemble complex choreographies by defining foundation elements, and relating new elements to the foundations.

In a museum exhibition setting, a curator leads a group of museum-goers around four walls of electronic imagery. In the traditional classification of the Tibetan pantheon, families of deities are classified by means of direction. Like the traditional ritual circumambulation of groups of people singing and dancing around a stupa, museum-goers circumambulate the walls (re-creating the ritual of invoking the deities of the different directions). The curator summons images appropriate to the symbolic meaning of each wall. Imagery pertaining to each direction can be stored on the corresponding wall. The curator explains the pantheon and the classification, the museum-goers gain an understanding of the pantheon, while at the same time, having ritually traversed around the universe.

In the ceremonial setting, video monitors are arranged in the form of a mandala, a cosmic diagram of the universe. An initiation ceremony is taking place on the raised podium of the mandala: the entrance of a new student into the world of Thangka painting. A Thangka master guides the students' first painting strokes in the ritual. At various points in the ritual, and depending upon time, season, astrological or ecological conditions, presiding deities and tutelary deities are summoned, and their presence is marked on the video monitors. On this occasion, through the repetition of the ritual, the community of teachers and students is regenerated.

These scenarios elucidate a variety of possibilities for the integration of electronic image in cultural settings. Common to all of the scenarios is the notion of performance: chanting, singing, dancing, impersonating divine beings, acting out stories, presenting primordial time in contemporary time, isolating and preparing special places and times for the activities (Schechner 1977, pp. 197–224). As in all theatrical traditions, these scenarios transform space into a special place.

Efficacy of the ritual is contained in the enacting of the script, whether the script is a proportioning diagram for painting, words for chants, or patterns for circumambulation. For the participants, doing is as important as the resulting communication. 'We overlook that while we may believe that the anticipatory rite has no physical effect in the right direction, the rite itself is a formal expression of a will directed to this end, and that this will, released by the performance of the rite, is also an effective force, by which the environment in its totality must be to some extent affected' (Coomaraswamy 1977).

Electronic Thangka in a museum

The examples above are intended to be evocative and to illustrate some possibilities of integrating new media with traditional rituals. These settings are capital intensive, and we did not pay attention to the economic aspects of integrating the electronic technologies into the cultural setting. Switching to the practical plane of contemporary life, and translating our visions to an economically feasible possibility, we chose to implement a museum 'sketch book' aimed at the education of a general museum-going public who have had little exposure to Tibetan art. While the broad vision of the project is the use of electronic technologies in the live transmission in the original craft setting, the museum setting offers an opportunity to implement the vision in a modified craft setting. In a museum setting, the modified craft community consists of museum management, philanthropists, art historians, and the museum-going public. A small portion of the museum-going public is familiar with the Buddhist practices; the great majority is unfamiliar with Thangka painting, or perhaps even with Asian art. For the museum-going audience, education is combined with entertainment. Due to limitations of space, museum galleries display only a small number of objects from the permanent collections. Objects on display are typically labeled with textual descriptions indicating date and place of creation, style and size. Museum educational tours and special workshops generally concentrate more on communicating particular visual facts about art objects, than about the universal ideas embodied by them.

For the museum gallery setting, we built an interactive gallery sketch book installation by which museum-goers can gain a glimpse of the world of Thangka painting. Particular to the museum audience, we faced particular challenges which shaped the form and content of the sketch book: for the museum audience, the sketch book must be easy to use; its user interface model must be simple to understand; it must allow museum-goers open-ended explorations. With respect to the sketch book's construction, the presentation materials had to be authored by the museum curators who have little exposure to computers.

The Asian Art Museum of San Francisco realizes that special catalogues and traditional exhibitions do not convey to museum-goers the rich cultural context that is lost when an Asian art object, detached from its original environment, is placed in a Western museum setting. Absent in traditional museum exhibitions and catalogues are the artistic process and the cultural setting in which Thangkas are produced and used. The prime educational goal of the sketch book addresses this disparity by bringing process and cultural context of Thangka painting to the foreground of the museum-goers' experience.

With respect to the study and appreciation of Thangka painting, the sketch book's curatorial content aims at presenting a new way for the museum-goers to 'see' Thangkas, understand the elements of Thangkas, and understand the deep structure of Thangkas in their cultural context—the intertwining of mythologies, symbolism, tools, and processes of Thangka painting. Museum-goers can expect to appreciate the Thangkas, their compositional elements, and the inter-relationships between the Thangka elements, symbolic meaning of the different painting elements, and the use of Thangka in the rich cultural life of Tibet. For each deity in a Thangka, museum-goers can learn how to understand the deities' visual language: postures, gestures, landscape elements, thrones, ornamentation, etc. While the audience can see the inter-relationships of one Thangka element with another, they can, at the same time, appreciate how one painting differs from another.

Emulating the organization of book media, contemporary interactive video systems allow users to explore presentation material that is arranged in chapters and subsections, as hierarchies of predefined choices (Fox 1989). In contrast, for the electronic sketch book, an important goal is to allow museum-goers open-ended explorations, without forcing them into hierarchies and modes. Museum-goers must, at their own pace, control the educational experience with the sketch book.

From the user interface point of view, learning how to use the sketch book must be extremely simple. To avoid crowding problems in the gallery, optimal interaction time was set at eight minutes (Hoptman 1989). In the short period of interaction with the sketch book, the museum-goer receives an introduction to Thangka painting, learns how to use the sketch book, and explores Thangka related topics. Given the visual nature of Thangka imagery, another important goal is to preserve visual means to interact with the sketch book, keeping text-based interaction to a minimum.

Video database

Verbal models

Verbal models provide the Thangka painter with a description of the basic iconography of a deity. Verbal models are descriptive invocations known in Sanskrit as the Dhyana of a deity. They were produced by intensive concentrations of theologians and have been transmitted through recitations of verse. They are vivid, graphic, precise, detailed, and perhaps evoke more fantasy than the paintings. For example, consider the following Dhyana, which instructs the painter on the rendering of the goddess Sarasvati:

> Surrounded by delectable herbs of a Mt. Meru Grove, within a white and pure ocean of milk, she (Sarasvati) is seated on a white lotus with large petals, one face, two arms, her face calm, smiling, and lovely like a charming youth of sixteen years, her breast firm and high, narrow waist, in squatting posture; with her hands holding an instrument of many strings,…

For the sketch book, Dhyanas were recorded in conversation with the Thangka master and the monk. Dhyanas provide descriptions of the expressions on the faces of the various divinities, including metrics and nuances of rage and tranquility, the attire of deities, the settings in which they are depicted, their residences, vehicles, thrones, weapons and other possessions.

In addition to the classical Dhyanas, which provide records of traditional iconography, we recorded curators' contemporary interpretations of Tibetan art, critiques of paintings, and examples of deity identification, materials used in Thangkas such as stretching of the canvas, preparation of paints, etc. The painter also explained the relationship of the painting elements to ecology and the cosmos. For example, landscape elements symbolize the five elements: like the planet Earth, a Thangka needs to be balanced with the elements. In the recording of conversations, a goal was to balance the relationship between explicit and implied meaning. In the end the painter offered a ritual for the dedication of merit for himself, Xerox, the Asian Art Museum of San Francisco, the rest of the project's participants, and the project's positive aspirations to the world at large.

Visual models

The museum is especially fortunate in having a striking collection of rare Tibetan paintings. Over 800 images of museum Thangkas were selected to provide examples of deities' visual language. For each Thangka painting, images of the important deities, their visual features, close-ups of postures, gestures, and landscape elements, halos, thrones, etc. were recorded. An important benefit of recording images of the painting collection on video is that museum audiences can view the entire

Thangka collection instead of the few Thangkas on display. Since Thangkas are primarily vertical in proportion, and television monitors are more square, we lost resolution of the paintings for overall shots. However, video emphasizes minute details of paintings that would commonly go unnoticed by museum-goers.

The images of painting were supplemented by images of Tibetan landscape, architecture, people, sounds, ceremonies. Scenes of Tibetan imagery, such as temple architecture, rituals, landscape, flora and fauna, etc. serve as references to paintings. In traditional educational settings, the visual models consist of prototypical sketches and proportioning diagrams of the various deities. Figure 20.7 illustrates a proportioning diagram of the Goddess Tara. Thangka masters use these models as aids to instruction so that their pupils, in faithfully re-drawing and rendering the diagrams and sketches, acquire proportioning, compositional, and colouring skills.

The proportioning diagrams of Thangka masters are set forth in iconometric theory which, according to Gerasimova (1978), is the 'grammar of drawing, the science of mathematical proportions which imparts harmony to an image.' The Thangka iconometry is based on the proportions of the human body. Hence the measures of man—face, palms, and fingers—were its units of measurement. For the sketch book, proportioning diagrams of important deities were recorded as still frames. Next, we recorded catalogues of various gestures, landscape elements, shading techniques, and alternatives for Thangka elements.

The recordings of a master's drawing process preserves compositional sequence: generating a proportioning diagram, outlining within the diagram the deity's outline, and clothing and rendering the deities. The preservation of the painting sequence is most important for the painter, because the sequence is related to mythological beliefs, and the process of drawing is considered as an inviolable ritual. For example, the last strokes in the depiction of the Buddha's head are the outlining of the eyes. Just as the master describes, the final strokes of outlining the Buddha's eyes brings the deity to life.

Finally we recorded a short introduction on Thangka painting and an overview of the project, a video tutorial on how to operate the sketch book, and an audio dictionary of deity names and pronunciations.

Benefits of video

Video records, when played back under computer control, become a medium that well supports the recording of processes. Unlike static diagrams and drawings, the sketch book collects action sequences in which the Thangka painter creates images of deities. Because the museum-goers can replay the scenes and experience them, almost as if they were present with the master, they can understand the sense of time of the craft and its actions.

The video records capture the presence of the master. In recordings of the master's drawing process, the master would sing folk melodies and prayers invoking deities, and thus the recordings capture the master's presentation style, countenance and demeanour. The medium provides the ability to connect process-related records to the master's personality. Video unites the message and the messenger, the story and the story teller as an inseparable totality: the vehicle is as important as the destination. Live craft-persons author the presentation of their craft, and thus viewers can trace the origins of ideas and artefact to human sources. In contrast to the outwardly regimented appearance of Thangkas, museum-goers, in replaying the video records of the painting process, become surprised to see the joy and the relaxed nature of the artists who produce Thangkas.

The video medium allows curators to tie visual elements to the universal and timeless truths behind the visual elements of Thangkas. This differs from the traditional art historical focus on particular facts: style, size of the painting, donors, authorship. Video allows us to intermix, and uniformly present representations originally in diverse media. Combined this way, images and sounds of Thangkas, drawing processes, personalities of Thangka practitioners, cultural context of Tibet provide museum-goers with a rich source of cultural records.

In the replay of any particular segment, video preserves the engagement of the painting experience, more than just the residual marks or the songs, prayers, and chants. Tibetan spiritual practitioners believe that even though the Thangka is consecrated with special power, reproductions of the image may not convey the original spiritual presence of the image. The performance of painting deities is sometimes related to various times of the day, seasons, or ecological conditions. Thus the records of a deity painted in the morning, for example, might not sparkle in the viewer's mind with the same intensity when replayed at some other time. The recording of the painting process often takes at least a day's preparation for simple drawings; in editing and trimming these segments down to segments of less than seven minutes, we faced the challenge of preserving the essence, and ensured this by including the Thangka master in the editing process.

User interface

The user inter&ce model (for further details see Makkuni 1989b, pp. 339–369; Makkuni 1989a; Kramer & Makkuni 1989) is structured by clustering and inter-relating elements of the database. Using visual representations of a Thangka, the user interface provides the user with both a point of entry as well as a conceptual map of the database. Users, by interacting with the visual representations displayed on the computer screen, gain access into the video records.

Figure 20.7 Proportioning Diagram of Goddess Green Tara.

Diagrammatic interface

Thangka iconometry is based on the study of the proportions of the human body. Hence, the measures of man—face, palms and fingers—were its units of measurement. Although the deities of the pantheon are classified semantically as various emanations of the Buddhist essence (as we have noted earlier, represented by the different cosmic directions), compositionally they are divided into three basic classes: peaceful divinities, wrathful divinities, and ordinary human beings. Thangka iconometry relates the appropriate proportioning of the figures in accordance with this classification. In Thangka iconometry, face and palms are termed the ‘great’ or ‘large’ measures; they in turn contain 12 or 12.5 ‘small’ measures. Classification is based on the relationship between the overall length of a body of a portrayed divinity and the large measures. For example, in the class of peaceful deities such as the Buddha, the body is equal to 10 large measures, and a large measure is equal to 12.5 small measures (length of a hand, or face). In the class of wrathful deities, the body is equal to 8 large measures, and a large measure is equal to 12 small measures.

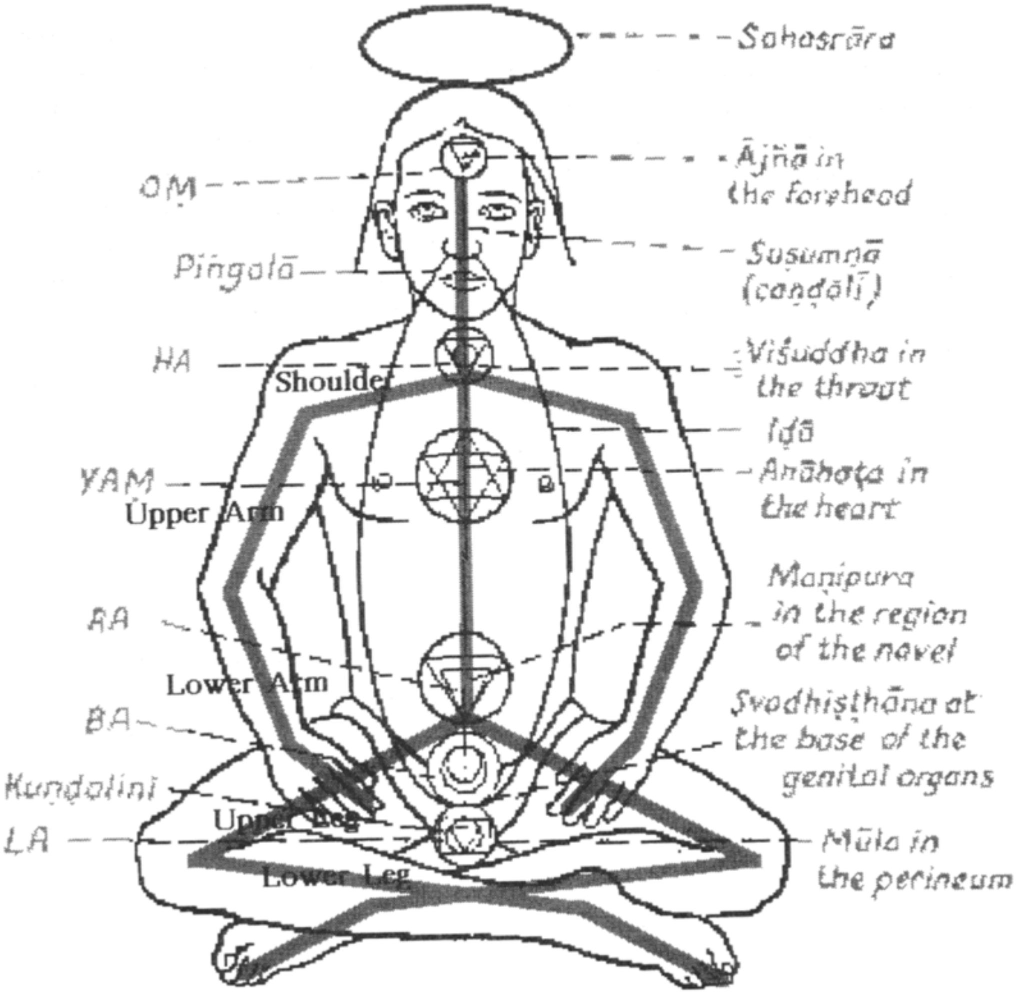


Figure 20.8 Extraction of a diagram from a yogic pose.

Derived from the iconometry, deities can be specified by means of a diagrammatic language: a collection of diagrams that represent various deities' essential graphic characteristics. Deities are modeled after ritual yogic postures. Figure 20.8 illustrates how the deity's yogic pose contributes to the construction of the diagram. In this figure, the spinal column is formed by joining various psychic centers, the 'chakras', on a yogi's body. These points on the spinal column serve as points for the construction of the centre of axis of a deity's diagram. A deity's spinal column may be vertical as in the case of Buddha, or may lean to the right or left as in the case of Goddesses. The spinal column is joined to the deity's arms and legs. Similarly, the arms and legs have varying positions. Arms may be raised, may stretch outward, or may relax in the lap, etc. Legs may be in locked, standing, or dancing positions, etc. Gestures, the varying positions of the deity's hands, are identified by the direction in which they point, and based on the position of the connecting arm. When classified in this way, the diagrams can be used for computer recognition or pattern matching.

A deity's diagram is a pictorial representation (see Fig. 20.9) which includes the dimensions of body parts, and proportioning relationships among parts. Based on the positioning relationships of body parts, a deity's yogic posture including hand gestures may be extracted from the diagram. Proportioning relationships also identify the deity's iconometric class. Hence, a diagram is a concise representation that not only serves as a graphic index to an image of a particular deity, but also provides clues about a deity's iconometric class. It provides the means to interact as well as index and retrieve elements from the database.

The diagrams may be considered as frozen steps in the dance movement of deities, like the dance of Shiva in ceaseless time (Vatsyayan 1984, pp. 191–201). Individual deities are specific instances of the prototype dancer, characterized by the positioning relationships of the limbs to the spinal column. In the dance tradition, the navel marks the centre of the dancer, which is constantly at rest. The different positions of the arms and limbs with respect to the centre, and the spinal axis, make up a kinetic vocabulary.

The museum-goer or beginning student manipulates the diagram by rotating the parts to change the positions of the deity's parts, or by changing the dimensions of the parts. For example, the student painter might begin with the lotus position

Figure 20.9 Diagrammatic indexing of deities.

(‘dhyana-asana’), subsequently select the deity’s left arm and gesture, re-position the left arm to point towards the knee, and change the left gesture from the gesture of equipoise (pointing to the right) to the gesture of earth pressing (pointing vertically down). In doing so, the painter transforms the lotus posture into the diamond posture (‘vajra-asana’). Positions of body parts can be changed by gesturing over a prototype diagram, progressively varying the positions of a deity’s parts. The modified posture diagram is matched against the dictionary of pre-defined diagrams by comparing its body parts with those of the diagrams in the dictionary. The closest match identifies the deity, and video records for that deity are laid out on an image of the matched deity.

The deity’s image is decomposed into regions, such as the region of the face, halo, nimbus clothing, jewelry, offerings, hand gestures, thrones, and landscape elements. The regions of the deity’s image are touch-sensitive for the painter’s interactive query. For example the student painter, by clicking at the region of the hands of a deity, obtains a menu of the titles of video records associated with that region. Video segments for the region of a deity’s hand include, depiction of that hand gesture, similar hand gestures of other deities, drawing exercises of that gesture, symbolism of the gesture, stories, etc. Similarly, clicking at the region of the face reveals the titles of the video records describing the deity’s face: the proportioning diagram for the face, verbal models for the eyes, stories, etc. Virtually every line on the deity’s image is available for interactive query. Selection from the menu of video segment titles executes the playback of the corresponding segment.

The cyclic process of manipulating diagrams, matching against deities, retrieving and viewing video segments, and manipulating diagrams once again, is how the user navigates and browses through the Thangka database. Interacting through diagrams has pedagogical value as a way of discovering the compositional essence of deities. The diagrammatic language provides the student with a taste of the compositional flavor, as well as glimpses of the content. It aids in the understanding of the visual language of the varying hand gestures and postures. As shown in Figure 20.9, we can imagine a diagrammatic interface implemented through the dancer’s postures, which can be useful in retrieving Thangka images of deities with similar dance postures.

However, this style of interface is too sophisticated for a general museum-goer to learn. The mechanics of operation of manipulating deity diagrams requires time to master. Also, the focus of the diagrammatic interface is on deities, and their visual aspects of deities. To retrieve information about landscape and cultural context, or the preparation of materials, becomes difficult to map on an interface that focuses mainly on the visual language of deities. This led us to design a simplified user interface model: the garden interface.

Garden interface

In the garden interface, the interactive learning experience of the museum-goer is like a walk through a garden of Thangka elements displayed on the computer screen —images of deities, their gestures, postures, thrones, auspicious elements, landscapes, flora, fauna, temple architecture (see Fig. 20.10). Museum-goers stroll through this garden using their fingers, stop at any particular area of interest, and by pointing at and touching that area to play back video records, obtain Thangka-related information for that area. For example, the museum-goer, by touching the region of the Boddhisattva Manjushree’s face obtains a choice of video records of Thangka painting related to the face: the proportioning diagrams of the face, invocations and prayers to Manjushree, a museum curator explaining the symbolism, or images of similar faces from the museum's Thangka collection. Selection of a video record executes the play back of the video segment on the corresponding monitors.

Museum-goers can, at their own pace, explore a variety of paths along the Thangka garden, and thus, control their educational experience. The garden of Thangka elements is not revealed at once; museum-goers discover the elements and paths as they begin to stroll through the garden. At the end of the interaction, the museum-goer receives a take-home souvenir: a hardcopy poster of the large garden with a plot of the museum-goer's path of discovery. The plot reveals the overall Thangka, and museum-goers discover the Thangka garden.

A typical state of a museum-goer's journey through the garden is shown in Figure 20.11. In this figure, the museum-goer has, in strolling through the garden, reached a particular image of the Buddha Sakyamuni in the garden of Thangka elements; subsequently the museum-goer touched the face of the Buddha. Upon touch, first, the face of the Buddha pictorially highlights with the sound of a drum. Second, the voice of the curator identifies the painting element to aid in the pronunciation, and provides a short phrase to describe the painting element. For example, in the case of the Buddha's face, the curator's voice first identifies the Buddha Sakyamuni, and second describes briefly his spiritual qualities and posture. Third, the face is connected by threads to three buttons appearing at the top of the drawing, each button corresponding to a video screen above the computer monitor. Each button displays a label that indicates the title of the relevant video material and its length of replay. The first frame of each video sequence is simultaneously displayed on the corresponding video monitor. By pressing these buttons the museum-goer can play back related video segments on the video database.

By touching various elements of the painting on the computer, the museum-goer gradually discovers the visual elements and compositional relationships in a Thangka painting. They begin to see the visual language of individual deities—postures, gestures, thrones, offerings, jewelry—and the variations of the language across other deities in the garden. By replaying the video they become transported into the world of the Thangka painter, the Tibetan monk, and in turn, Tibetan culture and the world of the deities.

This style of interface requires the user to relinquish preconceived expectations of Thangka instructional materials laid out in the traditional art history format. The museum-goer discovers the Thangka similar to the way a garden is experienced; at the beginning of a session they receive a short introduction on Thangka painting, and a tutorial explaining the mechanics of operation of using the user interface. At the end of the interaction a poster shows museum-goers their journey of discovery. This allows them to return and revisit areas in the Thangka garden that were previously unexplored.

Observations of use

The installation of the sketch book in the gallery setting (in October 1989) created a special performance space. Performance in the gallery includes a complex of events that take place from the time the museum-going audience is in the field of the installation until they leave. We made observations of museum-goers, how they form groups around the installation, and how they participate with fellow museum-goers and interact with the sketch book. We studied whether the system is accessible to use, and asked museum-goers about the messages they receive. These observations were carried out through paper questionnaires, open-ended video interviews, and computer records of interaction. From the questionnaires and interviews we came to understand the audience composition, their background, whether they had prior exposure to Tibetan art and electronic technologies, the time of interaction, their reactions to the instructional materials, the user interface and presentation of the exhibit, and the overall messages museum-goers receive. The computer records the history of interaction, beginning and endings of sessions, and maintains statistics of the various video segments that were summoned by museum-goers.

The composition of the audience ranged from ages 8 to 76, the average age ranging from 36 to 49 years. The majority of the audience had no prior exposure to Tibetan or Asian art, and little exposure to computers, but did have experience with television. Museum-goers, on the average, spend 10 minutes with the installation; a substantial number of them spend over 20 minutes; a few museum-goers spend over 2 hours. 50 museum-goers use the installation every day.

The introductions and tutorials did not assist the museum-goer to learn how to use the system. Instead, museum-goers learn by watching other visitors use the system. Since touching art objects is usually forbidden in the museum, museum-goers were initially reluctant to touch the installation. As soon as someone touches and becomes engaged in using the installation, a group of museum-goers form around the installation. Various roles occur among members in the group. One member in the group becomes the driver of the group, while the rest of the members offer suggestions and instruct the driver on how to navigate the Thangka garden. Members of the group of museum-goers seek assistance from each other, figure out how to use the system, explain insights in the content, and engage in discussions on Thangka related topics. Museum-goers welcomed the garden style interface and the use of multiple monitors.

At a purely compositional level Thangka painting is accessible to all, pleasing the museum-goer's eye with its vividness, and decorative charm. Deeper appreciation is usually reserved for the initiate of Buddhism, who has the ability to translate the

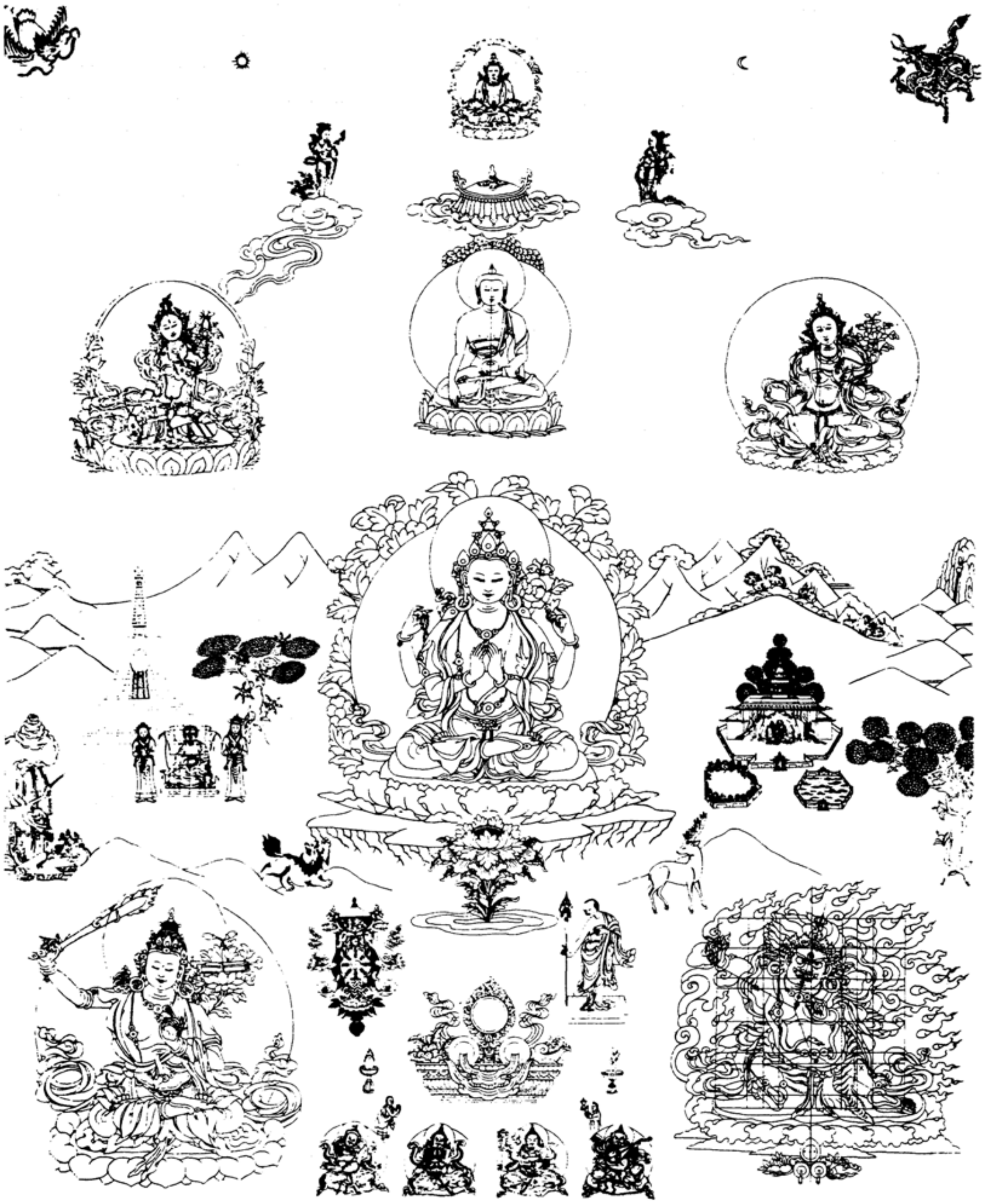


Figure 20.10 Garden of Thangka elements.

symbols into the religious experiences whose spirit the paintings seek to express. Museum-goers receive three main messages. First, a great majority of the museum-goers learnt about the compositional aspects and the visual language of Thangkas. The approach from composition informs the museum-goer about the Tibetan's pre-occupation with ornamentation, and the relationship of design and symbol. Museum-goers apply the new knowledge to the Thangkas in the adjoining Himalayan gallery, and learn to see Thangka composition in a new way. They discover details that would have, without the sketch book, gone unnoticed. Although museum-goers demand additional materials on Tibetan culture, they feel that the sketch book satisfactorily conveys the rich world of Thangka painting. Museum-goers, in particular, appreciated the presence of the Tibetan monk and master painter as authors of the instructional materials. They were positively surprised to see the relaxed and witty nature of the monk and the master painter despite the outwardly regimented appearance of Thangkas.

The second main message is about the positive uses of technology in the arts and life. Some museum-goers hold a view that electronic technologies are a hostile medium serving the worlds of finance and human exploitation, and are positively surprised to see their use in artistic preservation and dissemination. They especially appreciate the visual means of interacting with the sketch book.

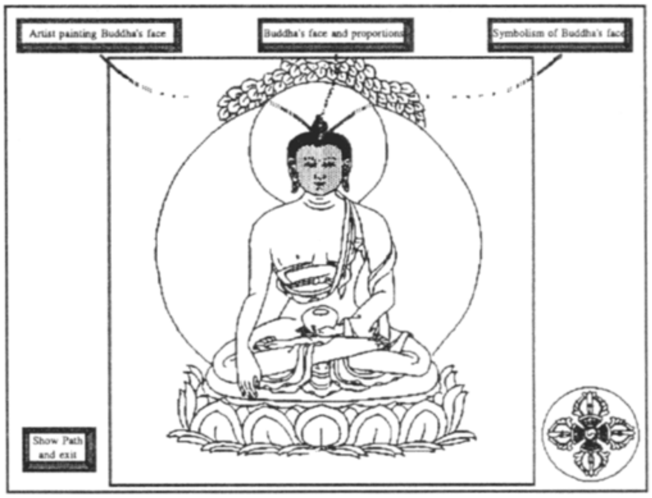


Figure 20.11 Interface to the Thangka Garden.

A few spiritually sensitive museum-goers receive a third message about the potential for the transmission of spiritual insight (Glessing *pers. comm.*). These museum-goers view the sketch book as a support in developing the mind and the spirit. Although the craft of Thangka painting is deteriorating, Tibetan art's system of symbols have a striking relevance outside Tibet. These museum-goers see the sketch book as vehicle for revealing psychological truths by making the symbolism more accessible to the day-to-day lives of a western audience. For example, the lotus reflects the concept of spiritual blossoming, of growing out of muddy waters to bloom above the water in full light, like an awakened being who has overcome darkness. Or the term 'Nirvana,' meaning the liberation from the cycle of re-birth, could be given contemporary relevance as freedom from day to day stress. Instead of using words such as 'deity', 'Gods', 'Goddesses', which emphasize the separation between the viewer and the pantheon, the re-design of the instructional materials makes it possible to bring people closer to the Tibetans' ideal of becoming awakened beings, developing their various virtues.

Conclusion

The electronic medium preserves the sense of 'traditional time' of the Thangka painter, uniting the message and the messenger, the painter's songs and the painting strokes, the painter's personality and the paintings. Living artists participate in the preservation and dissemination of their craft. Museum-goers trace the origins of ideas to human sources, and gain access to a different cultural point of view.

The present generation is very fortunate indeed to be exposed to diverse cultural points of view, which in turn, will contribute to future self-reflection. Electronic technologies can serve as vehicles of cross-cultural communication; although the electronic medium has a tendency to homogenize and standardize the presentations of culture, we hope to have illustrated an approach to the craft domain with a sensitivity towards capturing and disseminating the spirit of cultural operation behind the craft.

Acknowledgements

The project was supported by the generous contributions of The Xerox Corporation, The Asian Art Museum of San Francisco, The L.J. and Mary C.Skaggs Foundation, anonymous donors, Apple Computers, Inc. and ParcPlace Inc. John Seely Brown, Rand Castile, Steve Harrison, Dave Robson, and Bob Stults aided in conceptualizing the Electronic Sketch Book project. The design team for the museum implementation of the project were: Lama Wangyal, master Thangka painter from Dolpo; Tai Situpa XII, reincarnated regent of Tibetan Buddhism; Axel Kramer, computer scientist at Xerox PARC; and Terese Bartholomew, Molly Schardt and Richard Kohn, art historians and Tibetologists of the Asian Art Museum of San Francisco. Mark Chow, Brian Tramontanna, Theron Thompson, and Ed Foley were responsible for video production. Enrique Godreau, Steve Harrison, Gega Lama, Senge Lama, David Liebs, Dave Robson, Bob Stults, Karon Weber, and Frank Zdybel contributed to the development of an early prototype of the sketch book.

Austin Henderson, Candy Goodwin, Lucy Suchman, and Randy Smith offered valuable and thoughtful comments on the draft of this chapter. Figure 20.1, drawn by Jamyang is provided by Wisdom Publications, London. The diagrams on Figures 20.2 & 20.7 were drawn by Wangdrak, and are reproduced from Jackson & Jackson 1988.

References

Coomaraswamy, A.K. 1964. *The Crafts of India & Ceylon.* New York: Farrar, Strauss and Company.

Coomaraswamy, A.K. 1977. In *Traditional Art and Symbolism,* Lispey, R. (ed.), 71–178, Princeton: Princeton University Press.

Fox, E.A. 1989. The coming revolution in interactive video. *Communications of the ACM* **32 (7),** 794–801.

Gerasimova, K.M. 1978. Compositional structure in Tibetan iconography. *Tibet Journal,* **3,** 39–51.

Harrison, S. 1986. *Design and Media Spaces.* Palo Alto: Xerox PARC, (video).

Hoptman, G. 1989. Interactive video disk production. Presentation at the 1989 American Association of Museums meeting, New Orleans.

Jackson, D. & J.Jackson 1988. *Tibetan Thangka Painting: methods and materials.* Ithaca: Snow Lion Publications.

Kramer, A. & R.Makkuni 1989. Tools for constructing the electronic sketch book of Tibetan Thangka painting. In *Proceedings of the Conference on Technology of Object Oriented Programming Languages and Systems.* Paris: SOL.

Makkuni, R. 1987. A gestural representation of the process of composing Chinese temples. *IEEE Computer Graphics & Applications* **7(12),** 45–61.

Makkuni, R. 1989a. A Diagrammatic interface to a database of Thangka painting. In *Visual Database Systems: proceedings of the IFIP working conference on visual database systems.* Tokyo: North Holland.

Makkuni, R. 1989b. The electronic sketch book of Thangka painting. *Visual Computer* **5(4),** 227–242.

Pal, P. 1983. *The Art of Tibet: a catalogue of the Los Angeles Museum of Art collection.* Berkeley: University of California Press.

Pal, P. 1984a. *The Light of Asia.* Los Angeles: County Museum of Art.

Pal, P. 1984b. *Tibetan Paintings: a study of Thangkas Eleventh to Nineteenth Centuries.* Basel: Ravi Kumar & Sotheby Publications.

Schechner, R. 1977. *Essays in Performance Theory.* New York: Drama Book Specialists.

Smith, P.J. & E.Lucie-Smith 1986. *Craft Today: the poetry of the physical.* New York: American Craft Museum Weidenfeld & Nicolson.

Snellgrove, D. 1978. *The Image of the Buddha.* Tokyo: Kodansha International.

Stults, R. 1986a. *Media Space.* Palo Alto: Xerox PARC.

Stults, R. 1986b. *Representing the Process of Design.* Palo Alto: Xerox PARC (video).

Tucci, G. 1949. *Tibetan Painted Scrolls.* Rome: La Liberia Dello Stato.

Vatsyayan, K. 1984. Siva Natesa: cadence and Form. In *Discourses on Shiva,* Meister, M.W. (ed.). Philadelphia: University of Philadelphia Press.

Vatsyayan, K. 1986. The drama of ritual and the ritual of drama. In *Dimensions of Indian Art,* Chandra, L. (ed.). New Delhi: Agam Kala Prakashan.

Yanagi, S. 1972. *The Unknown Craftsman.* Tokyo: Kodansha International.

21

The implications of large scale image storage for primary archaeological research

ROGER MARTLEW

Introduction

Textual information in machine-readable form is now commonplace in archaeology, from excavation records and sites and monuments inventories to texts for publication produced on word-processors. The corresponding graphical information, however, is still confined to physical records in manual retrieval systems. The information portrayed by photographs, slides and drawings does not enjoy the same speed and efficiency.

The decreasing cost of computing power is, however, gradually bringing image handling down to a level which archaeological organizations can begin to contemplate. Some aspects of the technology are relatively new, and are still facing problems of development, let alone marketing. Other technologies have been around for decades, but have only recently been liberated by the availability of powerful, low-cost microcomputers to control them. Regardless of the development path of the delivery systems, the integration of text and graphics by computer will become increasingly affordable. Now is the time to address some of the issues surrounding the handling of images in archaeology, a discipline in which visual information is of vital importance (Martlew 1989). Unlike text, images offer a mode of communication which transcends language barriers. It is therefore appropriate to consider the storage, manipulation and transfer of visual data in archaeology at an international level.

The main issues reflect the extent to which text is necessary to allow a computer to provide efficient access to images. There are four potential levels of sophistication in image handling, starting with totally manual systems in which the images are stored in boxes or filing cabinets, and are described and indexed in a card catalogue. In the second level, the text index is transferred to computer, so that complex searches can be performed easily and quickly; the result, however, is a list of the physical locations in which the images are stored. The first real breakthrough in computerized image handling comes at the third level of sophistication, when the images themselves are transferred to some sort of computer-controlled medium. The result of searching the text index is that the image appears on a screen, either on a separate monitor or alongside the text of the catalogue. The final level of sophistication involves the replacement of text in the index by images, either wholly or in part. The images in the index may be lower-resolution or reduced copies of the images in the databank, or they may be generalised graphical representations (in either case they may be referred to as *icons*). An example of the use of such an index would be that, instead of asking the computer to match the string of characters in the word *handaxe* in the catalogue, the user would select an icon showing a graphical representation of a handaxe. This icon will have associated with it the locations in the databank of all relevant images. It will be apparent that while such a system may be easy to set up for a basic level of enquiry, and where there is a one-to-one correspondence between an icon and the concept which it represents, sophisticated interfaces will be necessary to deal with greater degrees of complexity in image content and retrieval criteria.

It was to investigate such issues that a project was set up in the Archaeology Department at Leicester University.[1] This chapter discusses the implications of the project for archaeological recording methods.

The *Archaeology Disc*: a pilot study

The main subject for the *Archaeology Disc* was British stone circles, with some material on related sites. This provided a reasonably well defined core of field monuments, presenting a range of morphological characteristics and a wide geographical spread. From the teaching point of view, stone circles also support an interesting variety of theoretical interpretations, such as astronomical alignments and geometrical construction, which often present difficult material for undergraduates to assimilate.

A visual record of such sites is important not only for showing the nature of the site itself, but also its setting in the landscape. The contrasts between sites such as Stonehenge and Castlerigg, or Balnuaran of Clava and The Hurlers can be seen instantly on the screen, without any need for lengthy textual descriptions. Given budget limitations it was to our advantage that finds from stone circles are relatively few, but they were included wherever possible. The videodisc therefore not only presents images of field monuments from the Shetland Isles to Cornwall, it also enables the user to see, simultaneously, a site and the artefacts recovered from it— artefacts which may in reality lie in a box in a museum store many miles from the site itself.

In considering the implications of the technology for the dissemination of visual information, some sense of scale can be achieved by comparing the videodisc with traditional printed media. Stone circles represent an excellent example of the rapid fall-off in the availability of visual information once the *classic* examples have been covered in, for example, general textbooks on prehistory. The same fall-off applies to other types of sites, and also to artefacts. In seven general books on prehistory, Stonehenge receives most coverage in half-tone plates; the twelve pictures are mostly black-and-white, with notable exceptions in a particularly glossy exhibition-related publication (Clarke, Cowie & Foxon 1985). The limited range of other sites shows the rapid fall-off in the publication of this visual information, from Avebury (8 plates) to Castlerigg and the Stones of Stenness (3 plates each). Four other sites each appear once, spread between three different books. These figures may not seem surprising in general books on prehistory, where there is a wide range of material to be covered. However, the main specialist book on the subject of stone circles (Burl 1976) contains only 36 black-and-white plates. Similar numbers of plates are presented in the corpus of Scottish chambered tombs (Henshall 1963; Henshall 1972), and major excavation reports provide visual information of this quality on a similar scale.

Despite the fact that visual information is so important in archaeology, the dissemination of such information is seriously restricted by the costs of traditional publication: few authors would turn down an opportunity to associate more pictures with their text, if cost were not a major consideration. The result is a significant limitation on the communication of knowledge and ideas in archaeology. Students can expect to see images of sites and artefacts from personal slide collections in their lectures, perhaps including unpublished material, but even this is not information to which they have easy access (Martlew 1990).

This is not to say that printed textbooks should be full of pictures, or should even attempt to present a comprehensive visual catalogue. Equally, the availability and portability of hardware still imposes restrictions on access to computer-controlled images. Paper- and computer-based media each have different strengths and weaknesses, and each is best suited to a specific role in the dissemination of information. An attempt was made on the *Archaeology Disc* to show at least one picture of every known stone circle, even if the site consisted of widely-spaced, low stones set among tall heather. Such images would never normally be considered for traditional publication, but they take up a negligible amount of space on a videodisc. The unit cost of putting an image onto videodisc is much lower than for traditional printing, but the enormous capacity of a videodisc makes production costs appear excessive. Over 2,500 still pictures on the *Archaeology Disc* are supplemented by eighteen minutes of full motion video, and almost all of the images are in full colour. The point of the comparison between textbooks and videodisc is to show the extent to which traditional printed media have failed to disseminate visual information, relating to sites, artefacts and the recording of excavations. An extremely efficient medium is now available not only to do this job, but also to increase the user’s power over visual data.

One particular strength of videodisc technology is its capacity for combining still pictures, moving images and sound. Archaeological material is not noticeably mobile or noisy, but there are ways in which the techniques of television can supplement the information about artefacts or sites recorded conventionally by still photography or drawing. Videodiscs, however, offer better access to the images than linear video or film, and it was important to locate relevant material for the *Archaeology Disc* which would allow us to investigate the implications of this for archaeological recording. Material on stone circles was located in a recent television programme, including rare (and genuine) footage of the winter solstitial sunrise at Newgrange in Ireland, a powerful way of introducing the subject of archaeoastronomy to undergraduates. Since there is full control of the two audio channels on the videodisc, the commentary, which is aimed at a lay audience, can be turned on or off. When producing new material for videodisc, the audio channels could be used to provide commentaries at two different levels, or in two different languages.

Even with this moving video, only about half of one side of the *Archaeology Disc* was used, so spare space was sold to anyone who wanted to try out the technology for a share of the production costs. This is a useful strategy for mitigating the cost of producing a videodisc in a small organization. Several archaeological bodies took the opportunity to test the technology with their own material, including the Royal Commission on the Historical Monuments of England (RCHME), the National Monuments Record of Scotland and York Archaeological Trust.

Optical discs for image archives in archaeology

The need for improved storage and retrieval systems for image data in archaeology is not hard to define. Videodiscs offer tremendous potential for applications ranging from the tens of thousands of excavation slides held by an urban unit, to the millions of slides, prints and drawings held in national collections such as the picture library of the RCHME. The inclusion of a range of material from York Archaeological Trust, the RCHME and the National Monuments Record of Scotland on the *Archaeology Disc* enabled the staff of these various institutions to see their own familiar images as portrayed by video. The result was a useful definition of the best role for interactive video in this archival context, and the identification of its limitations.

Visual catalogues

Videodiscs are already in use in some museums (most notably the Prins Henrik Maritime Museum in the Netherlands) as visual catalogues which can be searched by visiting scholars and members of the public. The videodisc serves as a quick and easy way of making a *first pass* through the available information. Instead of just getting a print-out listing the photographs or artefacts in a collection which match the user's interests, the user can actually browse through the pictures or see the artefacts straight away. This is particularly useful for archaeologists when trying to locate, for example, the best selection of aerial photographs of a group of crop-mark sites. As with text retrieval, locating the information by standard query on site name or location is relatively easy without a computer, since that type of query is anticipated by the archive's manual cataloguing system. The real advantages arise when the power of the computer to search on a number of different cataloguing fields simultaneously is combined with the almost instant retrieval of the image itself, not just its storage location.

This, however, is about as far as videodisc technology currently goes. One significant factor counts against videodiscs becoming the final archive for visual information, and that is the problem of image resolution. Analogue video screens contain 625 lines vertically in PAL format, with 768 square pixels on each line. At a distance of more than six times the height of the picture, the human eye cannot detect the difference between the lines (Clark 1987, p. 61). It is common for users of videodiscs, however, to be sitting closer to the screen than this in order to interact with the image via mouse, keyboard or touch-screen. The American NTSC format uses a lower 525-line resolution, and the two systems are incompatible. In order to show the whole of a large plan or artefact, small details have to be sacrificed. Close-ups have to be built in at the disc production stage, since any zooming of a digitised image from the finished disc will only show a close-up of the 625-line resolution picture. For the serious, in-depth study of photographs of crop-marks, finely decorated artefacts or detailed drawings, analogue video does not yet provide a solution for the professional. It is arguable, however, that at this level of study there will never be any substitute for the real thing (except in the case of ephemeral crop-marks), since ultimately photographs also have the same problem of resolution. What is important is the rapid and flexible identification of material for further study.

Heritage documentation

Videodiscs are being used in a number of projects to record standing buildings, works of art and other artefacts which come under the general heading of *cultural heritage*. A recent report identified 16 *cultural catalogues* on videodisc in use in museums and libraries throughout the European Community (Commission of the European Communities 1988, p. 259). A major project began in Italy in 1986, with the government spending 600 billion Lire on combined text and image databanks to record the country's cultural heritage. 16 out of 19 projects examined for the DOCMIX report are based on optical disc technology, with titles ranging from *La presenza ebraica in Italia* (an image databank of ancient Jewish artefacts on videodisc) to *Torri e complessi fortificati di Roma medioevale* (towers and fortifications of medieval Rome). Another project, *Verso Genova Medieval*, allows surrogate travel through medieval Genoa using analogue images on videodisc and computer-generated graphics (Commission of the European Communities 1988, p. 244).

The SIRIS project (Sistema Informative per la Riconstruzione dell'Insediamento Storico) aims to create an integrated text, cartographic and image databank of historical settlements in the region of Emilia-Romagna. What is in effect a multi-media geographical information system consists of a micro-VAX running INFORMIX under X-Windows, and controlling a videodisc player which displays images on a separate monitor. The project cost over one billion Lire, and employed 96 people for two years (SIRIS 1989).

In France, 22 regional centres have been set up to make heritage documentation more accessible to the public. The level of documentation varies from region to region, but there is generally a heavy reliance on microfilm. At least one region, Languedoc, is developing a videodisc application, in this case on medieval stained-glass windows (Centlivre, Toche & Riou forthcoming). Other French videodiscs include one on mosaics and a series of discs on objects in the Louvre.

In Denmark, videodiscs are being used to display 2,000 background maps on which the distribution of prehistoric sites and monuments can be plotted using overlaid computer graphics. A project at the National Museum in Copenhagen has stored 105,000 pictures of artefacts from the museum's collection on videodisc, for use by visitors to the museum (Larsen 1989).

The above description of these projects shows the scale of operations at which videodisc technology comes into its own. The sheer volume of material which can be stored on a videodisc can in itself be a disincentive, both to potential developers and to funding agencies. The transfer of thousands of images to videodisc is labour-intensive, and therefore expensive, but it comes nowhere near the labour and costs involved in cataloguing the images, and in developing support materials in the form of a user-friendly interface and documentation. Archival videodiscs demonstrate their full worth on a national and international level, with consequent implications for project funding.

The future for videodisc archives

The main consideration for the future development of videodisc applications is the dichotomy between current specifications of optical media. The potential rewards are great—a computer-controlled delivery system which has sufficient image resolution to replace photographic archives completely. On the analogue side, the laservision format for videodiscs may be superseded by the development of high definition television, while on the digital side the Compact Disc (CD) format appears to have inspired the War of the Acronyms: CD-I (Compact Disc Interactive) and DVI (Digital Video Interactive) are fighting for a commercial foothold, and in the meantime CDXA (Compact Disc Extended Architecture) offers a partial stage in the development of combined digital video, audio and data. Confusion over hardware standards, interfacing and software standards is adding to the inevitable claims and counter-claims in an industry which is still very young, as rival manufacturers battle for potentially lucrative domestic markets. Quoting the editor of *The Videodisc Monitor,* Frenkel (Frenkel 1989, p. 875) points out that there is still 'a long way to go before digital formats are anywhere near as cost effective or data dense for full motion video and large databases' as analogue videodiscs. In the meantime, there are many important issues about the indexing of image databanks, and user-friendly interaction with them, which must be explored (Clark 1987). Current technology is perfectly adequate for investigating these issues, and it would be short-sighted to defer this work in the expectation of technological changes some time over the next ten years. Now that there is a videodisc containing a range of material relevant to archaeology, many of the problems specific to this field can be tackled

If future work with optical discs—of either analogue or digital format—is led by the technology, laservision videodiscs will remain in a relatively small, specialised niche serving large-scale stills archives (such as picture libraries), and commercial training needs which require full screen, full motion video of broadcast quality. If, however, future work is applications-led, videodiscs will form a significant part of a multi-media environment alongside smaller scale archives (such as excavation archives), and ad hoc collections of mixed text and graphics on CD. Analogue video technology is well-developed, and available now. Applications requiring full-colour, full-screen moving and still images can be developed and disseminated using internationally agreed video and de facto interface standards. Digital image-handling techniques for optical discs are still developing, and are struggling to reach international markets and standards. If long term archiving is a requirement, with a need for independence from hardware developments, then material should be originated on, or copied to, photographic film, along with colour reference keys to monitor deterioration (Larsen, *pers. comm.*). This back-up can then be used to produce a digital optical disc when hardware, software and image resolution standards have been developed and adopted by an appropriate user base.

Conclusions

The integration of text and graphics in a computer-controlled environment offers tremendous potential for increasing the efficiency with which archaeological information is handled. The main issues fall into two groups: the inevitable uncertainties over the speed and direction of hardware and software development, and the problems of managing large amounts of visual information. Only time will present answers to the former, while the hardware which is available today can enable work to start on tackling the latter.

The extensive use of visual information makes archaeology an important area for development, not only in the context of national archives but also in the transfer of images at an international level. World-wide computer networks have shown what can be achieved when attention is paid to the translation requirements from one system to another. Videodisc technology is currently hampered by its reliance on analogue video standards, with different systems in America, Britain and Europe requiring expensive analogue conversion equipment. Manufacturers are, however, already beginning to solve this problem, with dual-standard players and interface boards which can accept either PAL or NTSC signals.

At a different level, collaboration is required on a world-wide scale to ensure the possibility of transferring images and image archives from one part of the world to another, irrespective of national languages. Translation systems will be required if there is a heavy reliance on text to provide indexing and navigational controls for visual information; if the power of images is used to overcome linguistic differences, however, a greater understanding of the semantics of images must be developed. A major breakthrough in the dissemination and processing of archaeological knowledge awaits such developments.

Notes

1 The project was funded by the University Grants Committee and the Computer Board for Research in Universities. Established under the *Computers in Teaching Initiative,* the project also investigated the use of interactive videodisc technology for teaching archaeology to undergraduates.

References

Burl, A. 1976. *The Stone Circles of the British Isles.* London: Yale University Press.

Centlivre, B., O.Toche & Y.-U. Riou forthcoming. *Architectural Heritage: new technologies in documentation.* Strasbourg: Council of Europe.

Clark, D.R. 1987. Twenty-first century books: an assessment of the role of videodisc in the next 25 years. In *Interactive Media,* Laurillard, D. (ed.), 60–73. Chichester: Ellis Horwood.

Clarke, D.V., T.G.Cowie & A.Foxon 1985. *Symbols of Power at the Time of Stonehenge.* Edinburgh: Her Majesty's Stationery Office.

Commission of the European Communities 1988. *DOCMIX State-of-the-Art and Market Requirements in Europe Electronic Image Banks.* Report EUR 11736 EN. Brussels-Luxembourg: Information Industries and Innovation: Directorate—General Telecommunications.

Frenkel, K.A. 1989. The next generation of interactive technologies. *Communications of the ACM* **32(7),** 872–882.

Henshall, A.S. 1963. *The Chambered Tombs of Scotland. Volume One.* Edinburgh: Edinburgh University Press.

Henshall, A.S. 1972. *The Chambered Tombs of Scotland. Volume Two.* Edinburgh: Edinburgh University Press.

Larsen, C.U. 1989. *Video disks and image processing.* London: Council of Europe & the Royal Commission on the Historical Monuments of England.

Martlew, R. 1989. Pictures and Pedagogy: teaching archaeology using image archives on videodisc. *Science and Archaeology* **31,** 12–19.

Martlew, R. 1990. Videodiscs and the politics of knowledge. In *Humanities and the Computer: new directions,* Miall, D. (ed.), 39–48. Oxford: Oxford University Press.

SIRIS 1989. SIRIS: Sistema Informativo per la Riconstruzione del'Insediamento Storico. *Informazioni* **Anno V, nuova serie, 1.** Bologna: Istituto Beni Culturali e Naturali della Regione Emilia Romagna.

22

The development of dynamic archaeological publications

SEBASTIAN RAHTZ, WENDY HALL & TIM ALLEN

Introduction

This chapter[1] discusses potential applications of multi-media technologies in archaeological and historical publications, and reports on experiences in building electronic books in these disciplines.

We start by discussing the specific question of electronic versions of archaeological publications, and in particular the design of excavation reports for the next decade. We distinguish between the 'passive' electronic archives which are already being built, and 'active' documents which will correspond to the personalized syntheses which are the traditional outcome of excavation. The second part of the chapter deals with an experimental system built from materials prepared for a 'normal' book, in order to emphasize the problem of the very large backlog of material which we will have to place in electronic form. Finally, we analyse the use of videodisc in dynamic or 'authored' applications, as opposed to simple 'encyclopedic' approaches, in the context of historical records. Our discussion is designed to remove an unfortunate perceived difference between educational material and research publication, by demonstrating how a common base can be used by a wide variety of readers.

In contrast with some other chapters in this volume, we are not concerned with a general model for presentation of archaeological arguments (eg. Stutt & Shennan 1992), or looking at strategies for archival data storage (eg. Wilcock 1992), but with suggestions for publication of data and interpretation.

Electronic excavation reports

Archaeology is both in a good position, and has a strong need, to take advantage of the opportunites offered by modern computer hardware, software and networking for the dissemination of information:

1. Much of the information is very specialized, and it is not economically viable to publish material in a conventional book which will only be read by a few hundred people.
2. The discipline at the field level is very data-intensive.
3. There is already considerable use made of graphical representations of the data (plans, drawings, photographs, stratigraphic diagrams), none of which are reproduced easily in a traditional book; experiments such as Rahtz & Hirst 1976 in using a landscape orientation for A4 pages, to better display plans which are commonly drawn in such an aspect ratio, have not been a success, and are unpopular with librarians.
4. Recognition of, and *discrimination by,* colour is an essential part of the archaeological record, but is seldom shown in publication. We stress this point of discrimination, since many applications of multi-media technology only utilize colour to allow *recognition* of a previously known fact. An obvious application in archaeology is the publication of ceramic thin-sections, where colour is an important *discriminant* between a new unknown object and an existing object in the database. Comparison between a picture of something and one's memory of it need not be precise; comparison between the colour of a ceramic fragment held in the hand, and a range of examples in a database, needs to be more precise.
5. Many archaeologists already have large quantities of machine-readable data, and considerable experience of formalizing information.

We should not, however, imagine that our problems are unique; most of the points cited above are common to all modern sciences (the contrast is greater with the arts disciplines with which archaeology is often associated.) We refer specifically here to the case of field work reports; the reader must bear in mind that the construction of these volumes will vary according to national practice and finance; our experiences are biased by the practice in Britain since the publication of the Frere Report (Frere 1975), which advocated a quadripartite division of 'information' into Level 1 (physical data), Level 2 (site records), Level 3 (archival interpretation) and Level 4 (synthesis). The question of an economically-important (ie. large enough) readership is important, but maybe the crucial issue is the immaturity of archaeology as a data-handling discipline. The enormous explosion in field work across the world in the last twenty years has not been accompanied by a comparable set of methodologies to compress the data to an understandable form, despite the long-term arguments of Jean-Claude Gardin and the logicist movement (Gardin, Guillaume, Herman, Hesnard, Lagrange, Renaud, & Zadora-Rio 1988). Even the tools we do have, such as the Harris Matrix to represent complex stratigraphy, have been difficult to turn into 'publication' (though we may note considerable interest in computerizing the matrix, such as the work in the UK (Ryan 1988; Alvey 1990; Huggett & Cooper 1991; Boast & Chapman 1991), France (Desachy & Djindjian 1991) and Germany (Herzog & Scollar 1991)). This has resulted in publications which consist on the one hand of undigested data and on the other hand of synthetic interpretations which are difficult to relate to the data elsewhere in the book. It is not difficult to pick up a typical excavation report, look at a plan, locate a feature, and be quite unable to find detailed information about it.

The fact that we are still developing new methodologies for data presentation means that we can adopt new facilities with greater ease than subjects which have invested heavily and successfully in existing methods.

Structuring excavation reports?

How can we best exploit the possibilities of electronic publication? A clear application (also discussed below in the context of historical documents) is in providing sophisticated interfaces to what we may characterize as *passive documents*, collections of data that impose no interpretation on the user. A typical example would be the use of videodisc and touch screen in the education centre of the York Archaeological Trust (Maytom & Torevell 1991), which offers an excellent learning environment, but does not actively promote an interpretation.

Do we wish to 'promote an interpretation'? The aim in producing a report is to allow the reader to learn from it, and we can distinguish between two possible models of learning:

- The cognitive approach, whereby the reader associates together atomic 'facts' to build up a picture of reality, and all the facts form a network of links and supports. Here, the author simply needs to have all the information available, and provide mechanisms for the reader to put together a private associative pattern;
- The procedural approach, in which the author marshals and *ranks* the information and presents it to the reader as a waiter delivers a meal, taking us from inference to inference, explicitly representing subordination and domination; this is the model of the traditional teacher, expounding a viewpoint.

It may seem that the former is more natural (though clearly certain well-defined teaching aims can be best achieved by the latter method), but there is little doubt that we are not at all accustomed to learning in this way, and find it very difficult to marshal large quantities of atomic facts with no preconceptions of relative importance. The great success of hypertext in recent years has built very heavily on the former theory, but we would suggest that this has been a simplistic approach which must be balanced by a continued use of interpreted knowledge (see Rahtz, Carr & Hall 1989 and Rahtz, Carr & Hall 1990, for further discussion).

If we adopt the 'network of facts' approach, our methodology is clear: we assemble all the information relating to an excavation, and allow the reader to enter at any point and explore outwards. Facilities for these are well provided in the current generation of hypertext tools (see Schneiderman & Kearsley 1989; Nielsen 1990, for straightforward introductions to this kind of work), such as the ubiquitous Hypercard on Apple Macintosh computers. If we adopt the 'directed learning' model, this form of learning is that followed by most books, and we can produce an electronic book which mimics the navigational tools found in printed publications (Table 22.1).

The traditional *interpretative* book relies heavily on a notion of hierarchy—a series of chapters breaking down into sections, thence into subsections, and paragraphs. Each is subordinate to the one above, but may contain references either to other elements of the same kind (eg. 'see chapter 4') or floating illustrative material (eg. 'see Figure 4'). What is very difficult to achieve is low-level cross-referencing; thus while we can give broad references to another chapter (while discussing a set of hut circles, we say eg.

Table 22.1 Navigation methods in a traditional book.

Book tool	*Theoretical effect*
Title page	top level summary
Contents page	browse tree
Book thickness	coarse navigation
Running heads	medium navigation tree
Section headings	fine navigation tree

Book tool	*Theoretical effect*
Fonts	data typing
Cross-references	non-hierarchical links
Index	alternative hierarchical view
Pictures, tables, lists	atomic node types
Glossaries & footnotes	hidden modules

‘see the ceramic data presented in chapter 5.3’), it is not easy to cross-reference every single context or object number. More importantly, the very common structures like tables and catalogues are static, and do not permit of reader-directed re-ordering to test an interpretation. The widely used spreadsheet programs permit just this level of dynamic interaction, by presenting changeable models based on one set of data, and it is this kind of facility we wish to integrate into the new book.

Books of the encyclopedia type clearly do not form hierarchies of chapters and sections. The issue here is one of granularity; within each entry in an encyclopedia, the writer may very well utilize the apparatus of sections, making the whole book the equivalent of a library. An interesting experiment was made in the publication of Hodder (1989) to imitate the non-linear nature of encyclopedias, by printing chapters of a multi-author thematic book in a random order, not even numbering chapters from the start. Unfortunately, the only effect is to make it difficult to find a desired chapter, and the fanciful experiment is unlikely to be repeated in future printed books; only if the book were genuinely an entity which *would* be read from beginning to end would the experiment have been of interest. More convincing is Nelson (1974) which contains two books, one starting from each end, printed on either side of the pages; it is possible here to read two volumes at once, switching between them by turning the physical book upside-down.

Preparing an electronic source

There is sufficient doubt about what the final form of a dynamic excavation report should look like that we would be well advised to start with a *source* document from which various versions can be generated. Most modern archaeologists with any sort of computing facilities at all will word-process their text, and the result is often not dissimilar to Figure 22.1, supposedly prepared for a traditional book, but in fact betraying:

1. A lack of understanding of typesetting resources, biased by the limitations of typewriter-like keyboards (eg. ‘×12’ when ‘×12’ is meant).
2. Arbitrary visual effects; the indentation before the first set of finds is greater than that before the second—whether or not this is supposed to convey information is ambiguous.
3. Insufficient statement of the underlying structure. The reader and the author both know which the section is, but the information is conveyed via a series of elaborate visual clues, distinctly unamenable to automatic processing.
4. Not enough attention to *re-useability* of the text for future researchers.

A more useful model that we have adopted is that of *generic markup*. This involves explicit tagging of each piece of information with its purpose, rather than instantiating that purpose with a visual effect such as white space or a change of typeface. Figure 22.2 shows how much information *can* be added even to this short paragraph, distinguishing all the important points. It is now possible for a computer to unambiguously extract the text of the description for context 552 without any interpretation of effects (the codes used are Standard Generalized Markup Language, SGML—see Goldfarb 1990 for further details). Figure 22.3 shows a possible instantiation of this markup on a typeset page, using type face variations and type size variations to distinguish the elements. The importance of this two-stage approach is twofold: on the one hand, we can vary the effect later by re-interpreting the markup; where we used bold before, now we can use extra vertical white space. On the other hand, we can take the same explicit source, and start to automatically generate versions designed to be read on screen.

An experimental hypertext book—the *Electric Rough Ground Farm*

Ted Nelson[2] coined the term *hypertext* to describe ‘non-sequential writing’, or works which could not be expressed as a simple linear sequence of their contents. Hypertext characteristics include:

Branching texts where the reader is presented with a choice between several paths to follow through the work.

Interconnected texts which make reference to other parts of themselves or parts of other, separate works.

Active texts which modify themselves according to some particular criterion, for example by recomputing one of its own tables of figures from the latest available stockmarket information.

B.BEAKER PERIOD PITS

1. Description of excavated features

552 Circular pit 0.68 m across and 0.12 m deep. Saucer-profile, filled with dark humic soil with small stones and a few charcoal flecks. Adjacent to this pit on the north-west side was an oval ?posthole 0.15 m east-west by 0.12 m north-south and 0.07 m deep.

Finds: Pottery (Fig. 9: P7, P8, P9, PlO, Pll, P12, P13, P14)

790 Circular pit 0.60 m across and 0.25 m deep. There were two fills: clean brown clayey silt (layer 790/l) overlying dark greyish-brown clayey silt with charcoal flecks.

Finds: Pottery (Fig. 9 : P15); Flint (x12 fragments mail comprising 1 scraper, 1 serrated blade, 9 unretouched flakes, 1 calcined lump); 1 fragment of crystaline sandstone or quartzite.

Figure 22.1 Some raw input by word-processing author.

```
<section>Beaker Period Pits</section> <subsection>Description of excavated features
</subsection>
<feature id=552> <description>
Circular pit <m>0.68</m> across and
Cm>0.12</m> deep. Saucer-profile,
filled with dark humic soil with small
stones and a few charcoal flecks.
Adjacent to this pit on the north-west
side was an oval ?posthole <m>0.15</m>
east-west by <m>0.12</m> north-south
and <m>0.07</m> deep. </description> <finds> Finds: Pottery <ref figid=9>
P7, P8, P9, P10, Pll, P12, P13,
P14</ref> </finds></feature>
```

Figure 22.2 A version of the raw text with explicit markup.

Beaker Period Pits

1. Descri!tion of excavated features

552 Circular pit *0.68 m* across and *0.12 m* deep. Saucer-profile, filled with dark humic soil with small stones and a few charcoal flecks. Adjacent to this pit on the. north-west side was an oval ?posthole *0.15 m* east-west by *0.22 m* north-south and *0.07 m* deep.

Finds: Pottery (Fig. 9 : P7, P8, P9, P10, P11, P12, P13, P14)

Figure 22.3 A typeset version of the marked-up text.

Multi-media texts which are composed of material from disparate information media (such as text, graphics, sound and video). These works are known more generally as ‘hypermedia’ documents.

The idea of hypertext does not, of course, presuppose computers—even without the ideas of Vannevar Bush (Bush 1945) novels have frequently had concurrent threads of action (branches) as well as flashbacks (local interconnections) and literary cross-references (external interconnections; for example, we cannot read David Lodge’s *Small World* without reference to a huge body of older literary forms)—but the computer has made hypertext much more practicable in three ways:

Presentation A reader may follow branching pathways and connections to other documents by selecting the appropriate material (usually with a mouse) and pressing a button.

Speed The computer may retrieve information several orders of magnitude faster than a human can turn pages in a book or fetch a new book from library shelves.

Information unity The computer acts as a control centre, allowing the many media available to it to be manipulated in a uniform fashion.

All of these features are present in an experimental ‘book’[3] called the *Electric Rough Ground Farm;* this is a set of computer files which, used with the appropriate software, offer the screen-based equivalent of the traditional book. It is based on a report (Allen, Darvill, Green & Jones 1991) prepared by one of the authors (Tim Allen) on excavations carried out mainly in the 1960s by Margaret Jones. It represents the integration of a large body of disparate data, and is typical of a mainstream excavation report.

Following the idea of generic markup outlined above, the whole report, prepared using Wordstar for manuscript submission for conventional publication, was converted by hand to have all the important elements explicitly tagged. The markup used was the typesetting language, comparable to the SGML illustrated in Figure 22.2. This had the advantage that a parallel publication route could be followed by typesetting the book in the traditional way according to the publisher’s conventions.

The typesetting software could also be used to format the text for the screen, and this created yet another version of the file, effectively back where we started with a formatted ASCII text; the difference was that this conversion was repeatable and controllable. The ASCII text was then read into software called *Toolbook*, running under Microsoft Windows 3 on a PC. *Toolbook* is directly comparable to Hypercard on an Apple Macintosh, permitting the designer to create an application with:

1. Scrolling fields of text on screens with user-defined backgrounds.
2. User-defined menus at the top of the screen to select actions.
3. Buttons or other areas of the screen sensitive to mouse clicks, which are associated by the user with appropriate commands.
4. Text selection and marking with the mouse.
5. Tagging of words in the text to set off actions when clicked. In the *Electric Rough Ground Farm*, this is used to trigger references to figures and tables.

The result of this work is shown in Figures 22.4–22.13 with appropriate commentary, to give some flavour of the electronic book interface we are suggesting may be appropriate. Much of the design is based on the traditional book, but we have tried to incorporate both new visual cueing facilities (such as colour) and to provide some of the facilities needed for the 'network' mode of learning discussed earlier. We are conscious that this electronic book, though it is complete and 'useable', falls far short in many ways of the printed volume, though it has its own unique properties, and our continuing work will concentrate on improving the readability of the product.

The potential of videodisc

The development of the *Electric Rough Ground Farm* has enabled us to gain some insights into the re-presentation of formal interpretive prose. Large bodies of archaeological data, however, do not lend themselves to site-based publication, but would benefit greatly from more visual data. Here we need an enabling hardware technology, currently the videodisc (see Martlew 1992 for a fuller discussion of the medium). The videodisc is moderately well-established, and there is a good deal of experience in putting together large collections of pictures; although the current technology is likely to become obsolete in the longer term, there is little reason to imagine that the purer digital technologies (see below) will affect the system designer very much. There is no doubt that one way to utilize the enormous storage capacity of videodiscs is as large pictorial encyclopedias. Most videodiscs used in education today are collections of both still images and moving film sequences that lend themselves to this sort of treatment. However, just as finding information from a paper-based encyclopedia palls after a short time, so does finding information from an interactive videodisc system. After the novelty has worn off, what are we left with? Merely a large collection of pictures that can be accessed very rapidly from a computer?

The term *interactive videodisc* (IV) has become very passé in recent years. The early promise of the technology has not been realised, mainly because it failed in the domestic market and so the costs never came down far enough to make it viable for use in education or research. However, with the dawning of a new age of multimedia computing—that is the integration of text, data, graphics, video and sound in a single, unified computing environment—we are able to create environments which combine the educational excitement of interactive video, and combine it with the power of text and data processing and the visual impact of high-resolution graphics. How quickly such systems become commonplace depends on a number of factors including the cost of computer memory and processing power, the rate of development of new technologies, such as digital video, and the availability of stimulating software environments for manipulating multi-media data. The software techniques embodied in the current hypertext systems have already made a great impact on education, and we can expect a similar impact in 'research' publishing.

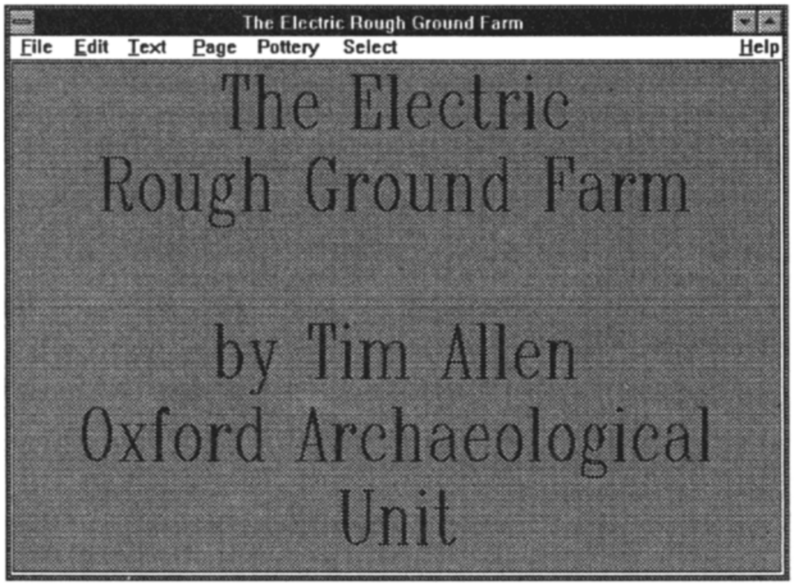
It is important to establish our visual identity at the start, and this opening screen serves to fix the screen size in the reader's eye, and the look of the menu at the top of the screen.

This welcoming message is displayed when the reader clicks on the WAC2 button in the lower right corner, as an example of temporary messages; this, like others in the book, simply disappears when clicked on.

This serves the same, very important, purpose as in a paper book, providing the reader with a visual map of what is to come; indentation is used to mark the hierarchy, and the scrolling page is programmed to jump to the appropriate section when any line is selected.

This shows the catalogue-type layout very typical of excavation reports; the main screen display consists of the large field of text, with a separate box above it showing the current section heading. The size, colour and typeface of this box vary according to the level in the hierarchy we are at (chapter, section, subsection etc). On the bottom, a scroll bar shows how far we are through the book, and can also be used to jump to places in the book. The buttons on the right hand side are used to (from the top):

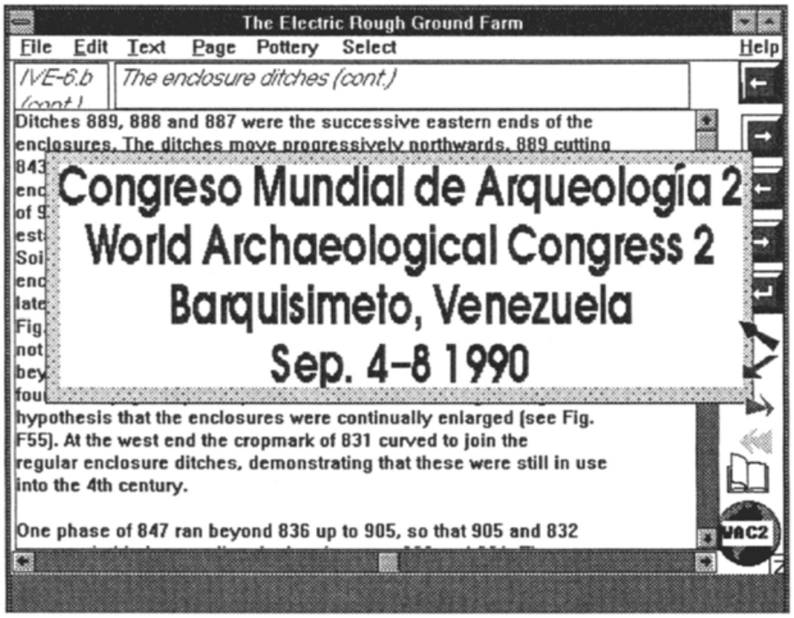
- *Jump to start of book.*



The Electric
Rough Ground Farm

by Tim Allen
Oxford Archaeological
Unit

Figure 22.4 Title page of the *Electric Rough Ground Farm*.



The enclosure ditches (cont.)

Ditches 889, 888 and 887 were the successive eastern ends of the enclosures. The ditches move progressively northwards, 889 cutting

Congreso Mundial de Arqueología 2
World Archaeological Congress 2
Barquisimeto, Venezuela
Sep. 4–8 1990

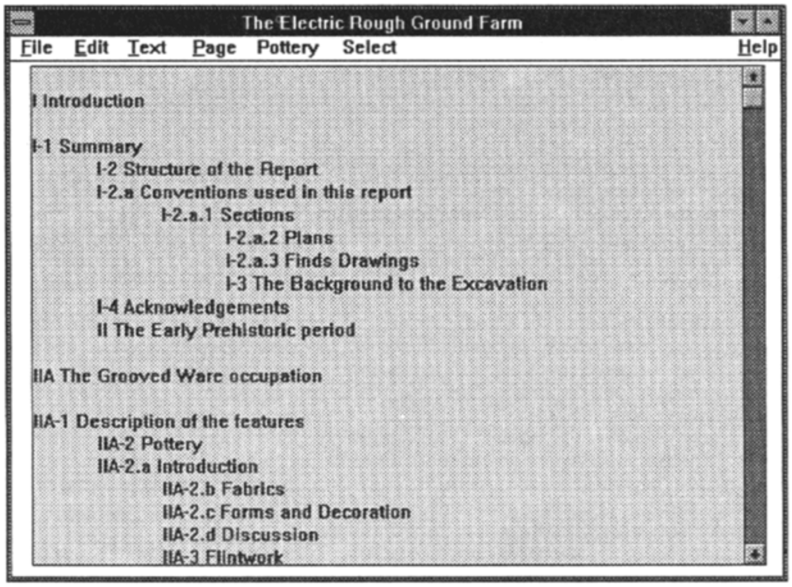
hypothesis that the enclosures were continually enlarged (see Fig. F55). At the west end the cropmark of 831 curved to join the regular enclosure ditches, demonstrating that these were still in use into the 4th century.

One phase of 847 ran beyond 836 up to 905, so that 905 and 832

Figure 22.5 The *Electric Rough Ground Farm* in Venezuela.

- *Go to next page.*
- *Go to previous page.*
- *Go to end of book.*
- *Go back to most recently visited page.*
- *Go up one level in the hierarchy; thus if we are reading a subsection, this takes us to the ‘parent’ section.*
- *Go down one level, to the next lowest element in the hierarchy.*
- *Go to the next comparable level in the structure tree; if we are at the start of a chapter, we jump to the next chapter, or if we are at the start of a section, we jump to the next section.*
- *We can return to the table of contents by pressing the ‘open book’ button.*

When a reference to a figure or table is encountered, the relevant information is fetched from another Toolbook *and displayed on the screen; seeing the figures is optional, and need not interfere with normal reading. Because of the sometimes complex typography, this table is displayed as a bitmap from the typesetter*



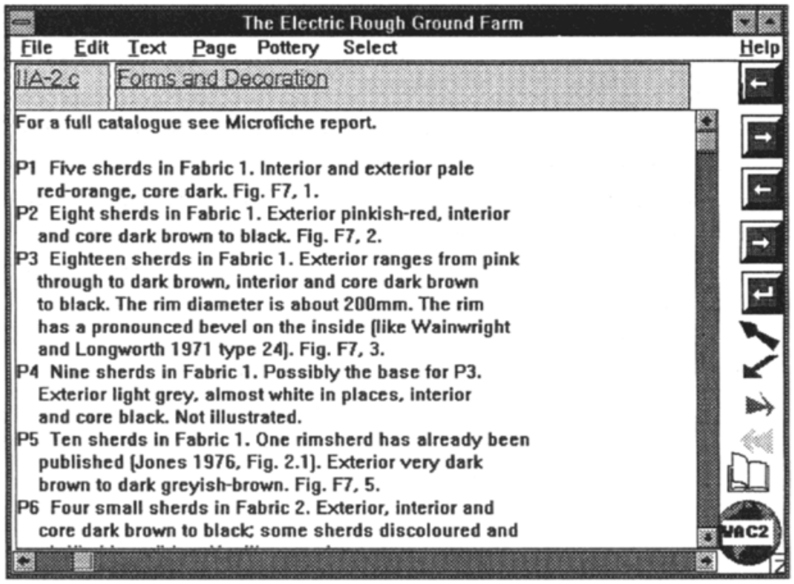
I Introduction

I-1 Summary
I-2 Structure of the Report
I-2.a Conventions used in this report
I-2.a.1 Sections
I-2.a.2 Plans
I-2.a.3 Finds Drawings
I-3 The Background to the Excavation
I-4 Acknowledgements
II The Early Prehistoric period

IIA The Grooved Ware occupation

IIA-1 Description of the features
IIA-2 Pottery
IIA-2.a Introduction
IIA-2.b Fabrics
IIA-2.c Forms and Decoration
IIA-2.d Discussion
IIA-3 Flintwork

Figure 22.6 Table of contents.



For a full catalogue see Microfiche report.

P1 Five sherds in Fabric 1. Interior and exterior pale red-orange, core dark. Fig. F7, 1.
P2 Eight sherds in Fabric 1. Exterior pinkish-red, interior and core dark brown to black. Fig. F7, 2.
P3 Eighteen sherds in Fabric 1. Exterior ranges from pink through to dark brown, interior and core dark brown to black. The rim diameter is about 200mm. The rim has a pronounced bevel on the inside (like Wainwright and Longworth 1971 type 24). Fig. F7, 3.
P4 Nine sherds in Fabric 1. Possibly the base for P3. Exterior light grey, almost white in places, interior and core black. Not illustrated.
P5 Ten sherds in Fabric 1. One rimsherd has already been published (Jones 1976, Fig. 2.1). Exterior very dark brown to dark greyish-brown. Fig. F7, 5.
P6 Four small sherds in Fabric 2. Exterior, interior and core dark brown to black; some sherds discoloured and

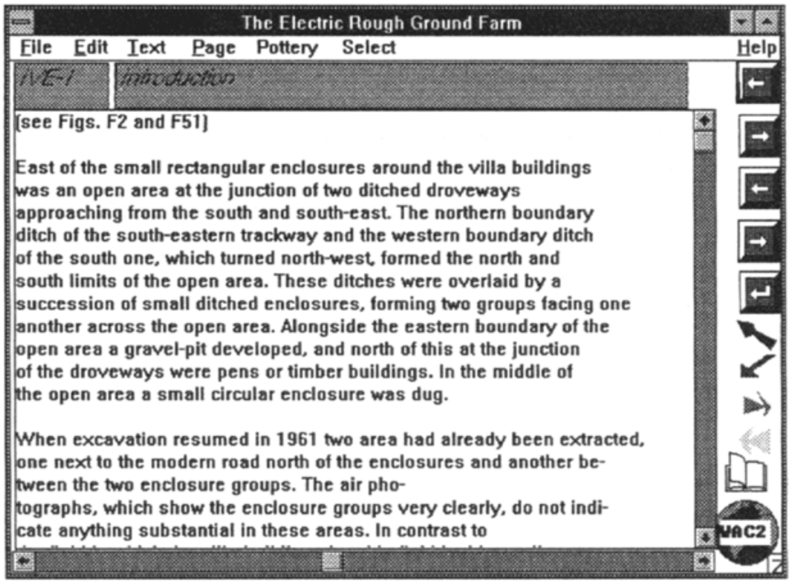
Figure 22.7 A typical page display.

The pottery details are displayed in a field superimposed temporarily over the main text; this information is extracted dynamically from a dBaseIII database.

From interactive video to multi-media computing

IV technology has been available since the late 1960s. As early as 1972, Nicholas Negroponte of the Media Lab at the Massachusetts Institute of Technology predicted the convergence of three previously unconnected fields of work—namely publishing, computing and broadcasting (Brand 1988). IV was seen as having an important role to play in these developments. However, until very recently, the cost of the technology made the use of video systems virtually non-existent outside of industrial applications. Most examples of videodisc material were in the industrial training sector, where it can be very cost effective, as well as educationally effective, to produce videodiscs and related software designed for specific training tasks. However, such materials are generally very expensive to make, and the industry training model provides no solution to the most effective way of exploiting IV in education (Hall 1988).

The development of new techniques such as surrogate walks demonstrated the enormous potential that IV systems could offer education. The combination of scope and power provided by the convergence of computing and video technology has



[see Figs. F2 and F51]

East of the small rectangular enclosures around the villa buildings was an open area at the junction of two ditched droveways approaching from the south and south-east. The northern boundary ditch of the south-eastern trackway and the western boundary ditch of the south one, which turned north-west, formed the north and south limits of the open area. These ditches were overlaid by a succession of small ditched enclosures, forming two groups facing one another across the open area. Alongside the eastern boundary of the open area a gravel-pit developed, and north of this at the junction of the droveways were pens or timber buildings. In the middle of the open area a small circular enclosure was dug.

When excavation resumed in 1961 two area had already been extracted, one next to the modern road north of the enclosures and another between the two enclosure groups. The air photographs, which show the enclosure groups very clearly, do not indicate anything substantial in these areas. In contrast to

Figure 22.8 Moving half-way through the book (note bottom scroll bar).

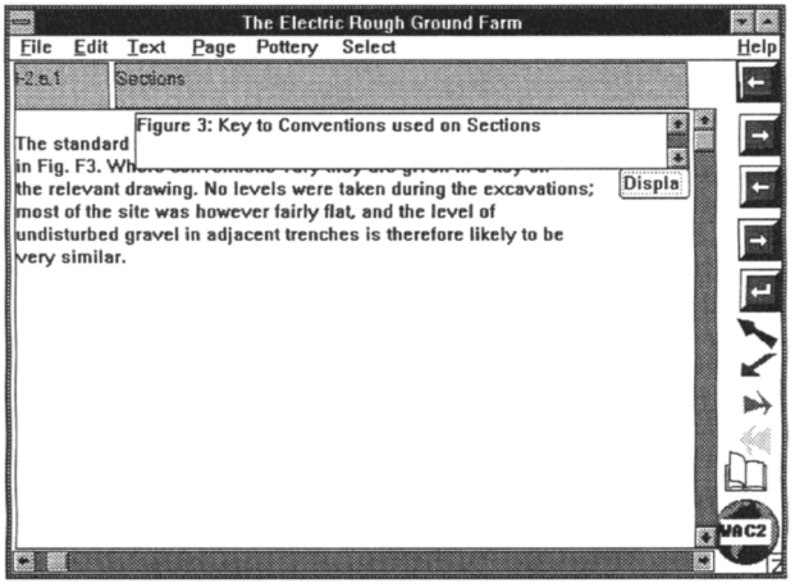


Figure 3: Key to Conventions used on Sections

The standard

in Fig. F3. Wh

the relevant drawing. No levels were taken during the excavations; most of the site was however fairly flat, and the level of undisturbed gravel in adjacent trenches is therefore likely to be very similar.

Figure 22.9 Summary display of Figure information.

the capability to greatly enhance and enrich the learning process. There are numerous areas of education and research where the only way to really understand or appreciate a concept is to see a picture or a moving sequence. The ability to access from a single workstation any one of thousands of still pictures, or moving sequence clips, offers the potential for new research tools and highly motivating learning environments. But the economics of the situation means that education cannot afford to create a new videodisc with specialized software for each teaching or learning task. A resource-based approach must be adopted, in which a videodisc contains a collection of material about a particular subject (which may be highly specialized or more general in its content depending on the nature of the project) which can be used for many different applications and by a wide range of users.

Even with the development of resource-based videodisc material, the penetration of videodisc technology into education at any level is still extremely limited. This has largely been due to the cost of producing discs, and to the fact that videodisc players are entirely stand-alone pieces of technology. The information on a videodisc cannot be shared or distributed in any economically viable way at the present time, and any user wishing to access the information from a videodisc must have a videodisc player and the appropriate videodisc available at their workstation. Generally, it is too expensive to equip whole laboratories of IV workstations, so the use of IV material is normally restricted to areas where research projects have provided the funds to set-up the equipment. But there is potential for radical change.

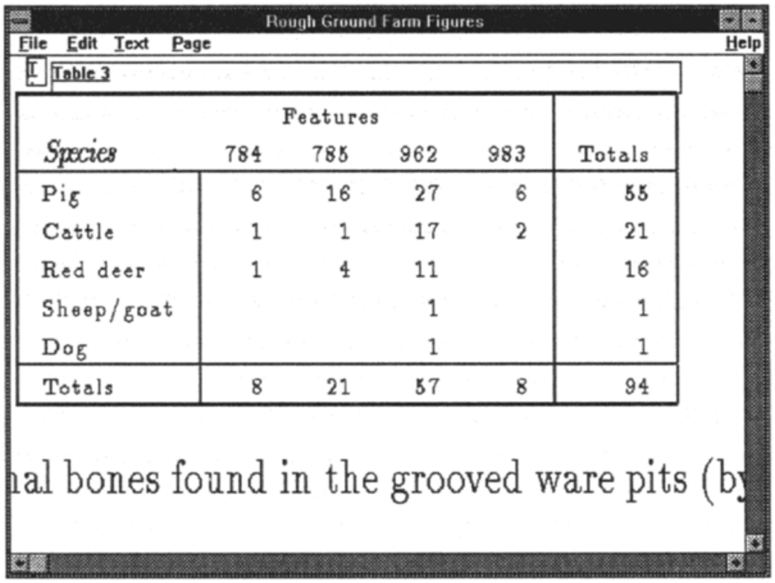
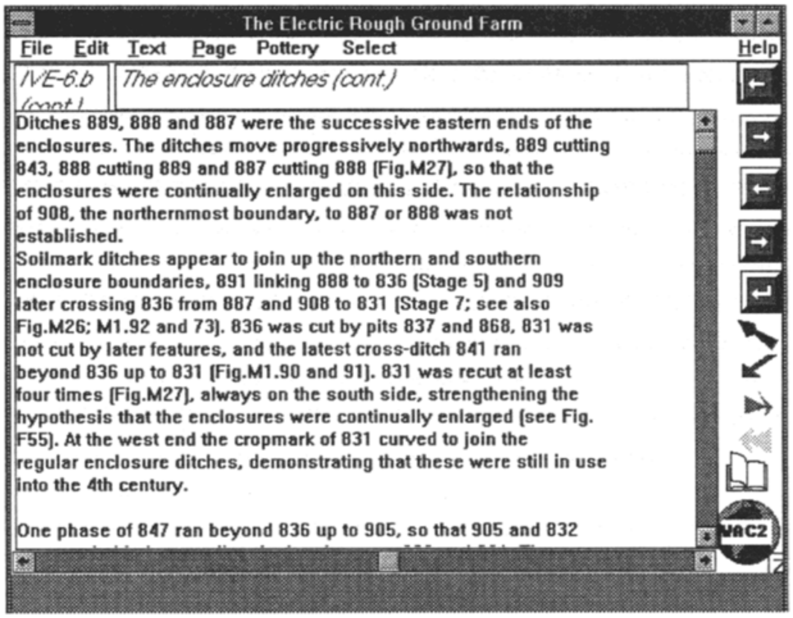


Table 3

Species	Features 784	785	962	983	Totals
Pig	6	16	27	6	55
Cattle	1	1	17	2	21
Red deer	1	4	11		16
Sheep/goat			1		1
Dog			1		1
Totals	8	21	57	8	94

al bones found in the grooved ware pits (b

Figure 22.10 Typeset table displayed as bitmap on demand.



IVE-6.b (cont.)

The enclosure ditches (cont.)

Ditches 889, 888 and 887 were the successive eastern ends of the enclosures. The ditches move progressively northwards, 889 cutting 843, 888 cutting 889 and 887 cutting 888 [Fig.M27], so that the enclosures were continually enlarged on this side. The relationship of 908, the northernmost boundary, to 887 or 888 was not established.

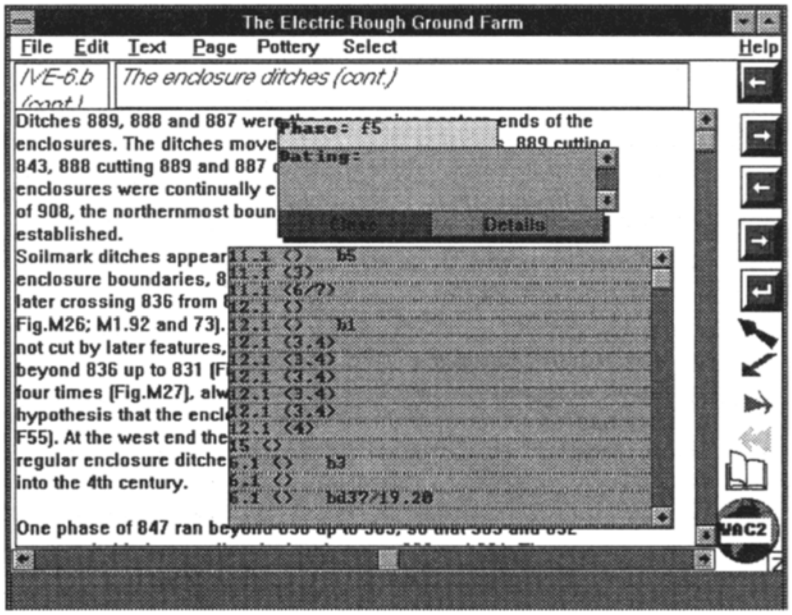
Soilmark ditches appear to join up the northern and southern enclosure boundaries, 891 linking 888 to 836 [Stage 5] and 909 later crossing 836 from 887 and 908 to 831 [Stage 7; see also Fig.M26; M1.92 and 73]. 836 was cut by pits 837 and 868, 831 was not cut by later features, and the latest cross-ditch 841 ran beyond 836 up to 831 [Fig.M1.90 and 91]. 831 was recut at least four times [Fig.M27], always on the south side, strengthening the hypothesis that the enclosures were continually enlarged [see Fig. F55]. At the west end the cropmark of 831 curved to join the regular enclosure ditches, demonstrating that these were still in use into the 4th century.

One phase of 847 ran beyond 836 up to 905, so that 905 and 832

Figure 22.11 Text describing contexts.

First, it is becoming cheaper to make videodiscs, especially where the material to be put on the disc is already available (ie. no new film has to be shot). This is due to the development of analogue WORM (write-once-read-many) technology, such as the Sony analogue WORM videodisc system. This enables the user to place film, video and other photographic material directly onto the videodisc, and is particularly useful for producing videodiscs to archive a collection of such material that is already in existence. Many academic groups have large collections of photographs, slides and video information for use in teaching or research that they can now, at a relatively low cost, transfer to videodisc. At the simplest level the resulting discs can be used, very effectively, in lectures and seminars, or can be used as the basis for sophisticated learning environments.

Second, and probably more importantly, we are on the verge of some important technological breakthroughs in the area of digital video. As previously stated, the main barrier to the use of IV is that it is an analogue technology. The information stored on analogue videodiscs cannot be passed around digital communications networks or stored on compact-disc (CD) technology. The latter has been an enormous success in the home market (in the guise of CD audio) and, as a consequence, the price of the technology has fallen dramatically since it was first launched. However, without the use of very specialised compression techniques, it is not possible to store any reasonable quantity of moving video on a compact-disc, nor to play it back on a computer screen. Hybrid technologies, such as VideoLogic's DVA (digital-video-architecture) 4000 board, enable



IVE-6.b (cont.)

The enclosure ditches (cont.)

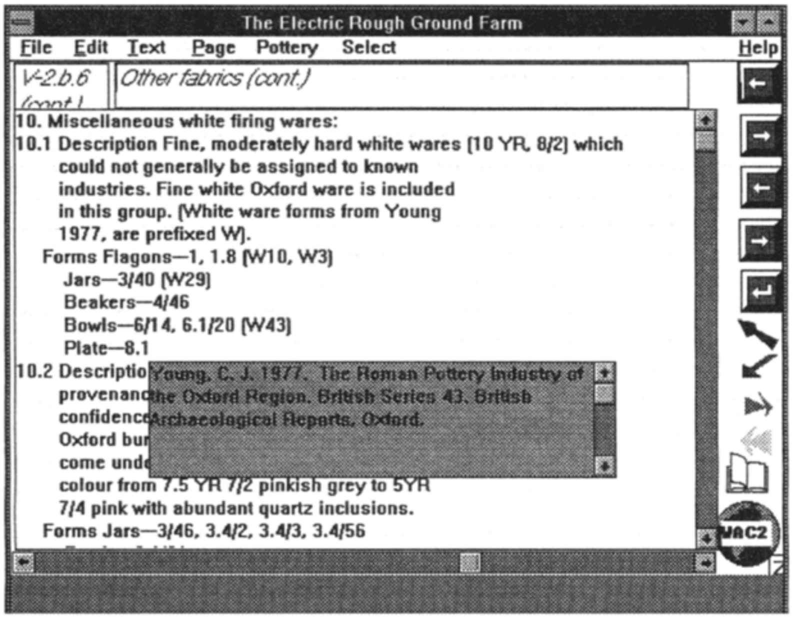
Ditches 889, 888 and 887 were … ends of the
enclosures. The ditches move… 889 cutting
843, 888 cutting 889 and 887 …
enclosures were continually e…
of 908, the northernmost boun…
established.
Soilmark ditches appear…
enclosure boundaries, 8…
later crossing 836 from …
Fig.M26; M1.92 and 73)…
not cut by later features, …
beyond 836 up to 831 (F…
four times (Fig.M27), alw…
hypothesis that the encl…
F55). At the west end the…
regular enclosure ditche…
into the 4th century.

One phase of 847 ran be…

```
Phase: f5
Dating:

11.1 <>    b5
11.1 <3>
11.1 <6/?>
12.1 <>
12.1 <>    b1
12.1 <3.4>
12.1 <3.4>
12.1 <3.4>
12.1 <3.4>
12.1 <3.4>
12.1 <4>
15 <>
6.1 <>    b3
6.1 <>
6.1 <>    bd37/19.20
```

Figure 22.12 Pottery details extracted from database and displayed for reader.



V-2.b.6 (cont.)

Other fabrics (cont.)

10. Miscellaneous white firing wares:
10.1 Description Fine, moderately hard white wares (10 YR, 8/2) which
could not generally be assigned to known
industries. Fine white Oxford ware is included
in this group. (White ware forms from Young
1977, are prefixed W).
Forms Flagons—1, 1.8 (W10, W3)
Jars—3/40 (W29)
Beakers—4/46
Bowls—6/14, 6.1/20 (W43)
Plate—8.1
10.2 Descriptio…
provenanc…
confidenc…
Oxford bur…
come und…
colour from 7.5 YR 7/2 pinkish grey to 5YR
7/4 pink with abundant quartz inclusions.
Forms Jars—3/46, 3.4/2, 3.4/3, 3.4/56

Young, C. J. 1977. The Roman Pottery Industry of the Oxford Region. British Series 43. British Archaeological Reports, Oxford.

Figure 22.13 Browsing pottery catalogue, with bibliographic citation displayed.

an effective compromise to be achieved, whereby the video information is stored on analogue videodiscs but is played back through the DVA board as a digital image on a VGA screen. But the promise of multimedia computing lies in the development of compression techniques that allow the development of full digital video technology and enable full-colour, moving video information to be stored on a chip within the computer, thus to be fully integrated with other types of digital data such as text, graphics and sound.

The two digital video technologies vying for pre-eminence are DVI (digital video interactive) and CD-I (compact-disc interactive); see Lock & Dallas 1990 for a discussion of their potential in archaeology. The former is being developed by Intel in association with IBM and the latter is the result of a joint venture between Philips and Sony. Both are based on compact-disc technology and both claim to be able to support full-colour, full-motion video. But with DVI the decompression of the video is carried out by a chip-set installed in conventional machines, whilst CD-I will initially be launched in the home market with a player that has its own built-in processor. We are in a transitional period from analogue to digital video, and it is unclear which technology will succeed and which fail. We can compare the situation with famous competitions of the past, such as that over the format of ciné film in the 1930s and the 1960s, the development of video tape technology in the 1980s, or the battle of the operating systems in the 1970s; these should teach us that the best technology does not always win. What is beyond doubt is that all the large computer manufacturers see multi-media computing as a major development of the 1990s. Just as the laser printer made desktop publishing a reality, so digital video will allow multi-media computing to move closer towards its real potential.

A three-level model for multi-media computing

The transition from analogue to digital video should not prevent us from exploring the potential use of video in multi-media computing environments, particularly in education. Building large pictorial databases, with the depth of coverage of subject matter that would be required for their effective use in a discipline, takes a long time. Any videodiscs we make today can be used as masters for digital video CDs in the future, so the financial investment and the time and effort spent on their production will not be wasted. Additionally, we can use them to develop effective software environments that exploit the enormous potential offered by the availability of large amounts of pictorial information at the personal computer workstation.

However, providing easy interactive access to such large amounts of pictorial information for both authors and readers is a task that will continue to occupy researchers. It is far more complicated to index a picture than a text document, and we need to provide environments in which the links between picture and text, and between picture and picture, are seamless to the user. Conventional access strategies for database systems were developed on the premise that the user knew what information was in the system and what questions they needed to ask to elicit the data they required. Full text-retrieval systems were designed to enable the user to search through large numbers of documents, usually on a key-word basis. The success of the search depended largely on the ‘correct’ choice of key-words and the expertise of the researcher. In order to effectively utilize large information systems in education and research, we must develop new access strategies that enable the learner to make associations between different pieces of information when they may have very little knowledge of the detail of the subject. There is a need for access strategies that build towards the synthesis of information from different storage media. Hypertext and hypermedia systems offer a possible solution to this problem.

The term ‘hypermedia’ is often used to describe a hypertext system that has been extended to incorporate other media, such as graphics, video or sound, in addition to text. There are a number of such systems under development both in research establishments and the commercial world, but there are as yet very few documented reports of their use in education. One example that has been widely reported is the Intermedia system (Yankelovich, Haan, Meyrowitz & Drucker 1988) at Brown University in the USA, which has been successfully used to create educational materials in a number of subject areas including English Literature and Biology. The ILAB project being undertaken at Southampton (Hall, Thorogood, Sprunt, Carr & Hutchings 1990) was established to evaluate the use of hypermedia as an interface to multi-media information systems in education. So far the interface has proved very effective as a means of providing students with access to multi-media data, particularly in linking textual information to graphical animations and video sequences. However, as would be expected, the main problem is that of disorientation within the program and we need to find better ways to help users navigate around large hypermedia systems. Maps, contents lists and indexes are all ways of easing this problem but they still leave the user free to roam around the information at will, which is not always helpful in a learning context. When users are unfamiliar with the subject matter, they need suggestions of what route to take next, either directly, or indirectly via a question. Directed paths, or routes through the hypermedia network of links, provide one means of achieving this. Tutorial modules can be included as elements in the hypertext database, access to which can be provided by links (commonly indicated by visual ‘buttons’ on the screen which can be selected by the reader) placed at appropriate points in the hypermedia system. The tutorial material can itself contain links back to the main database.

We therefore have a three-level model of access to a multi-media information system—the database level, the hypermedia level and the tutorial level. This allows the same information to be utilized in a multitude of different ways and by a wide range of users. The model is even more effective if the hypermedia links are stored separately from the content of the database, as it supports a resource-based approach to the development of multi-media computing in education and research. The software can be developed in stages, subsequent to the availability of the material in the multimedia database.

A case study: Microcosm and the Mountbatten Archive

Microcosm is a hypermedia system developed at the University of Southampton (Fountain, Hall, Heath & Davis 1990.) It has a powerful linking mechanism that allows for the creation of both specific and generic links in the documents system. A generic link is one constructed dynamically by the system. A reader may ask for more details about a particular subject, and instead of the author of the hypertext having to specify all the possible links in advance, the software attempts to find a reference automatically. The links are stored separately from the documents in the database (which may contain text, graphics, video or sound.) This separation allows links to apply not just to only one document, but to be used with a whole class of documents. This greatly reduces the authoring effort required, since a new document can be brought into the system and immediately allow linking. It also enables a network of information about a particular subject area to be created and applied to any documents brought into the system. Microcosm currently runs under Microsoft Windows version 3.0 and DOS 4.0. A Philips VP410 laser videodisc player is interfaced to the workstation to provide access to video information and the picture is displayed on a VGA monitor via a VideoLogic DVA-4000 board.

The Mountbatten Archive contains documents, photographs, films and audio tapes collected by Earl Mountbatten of Burma over his lifetime. The archive of approximately 250,000 text documents (letters, government papers etc.) and 50,000 photographs, contains information relating to all the important national and international events in which he played a rôle. The entire archive is kept and administered by the Hartley Library at the University of Southampton. Since this is a multi-media archive of some considerable size, the only effective way to make it available to researchers on an international basis is to make use of computer technology. The text documents can be scanned into a (full text retrieval) database and released, section by section, on CD-ROM as the cataloguing of different sections of the archive is completed. However, providing access to the photographs in the archive is much less straightforward. As part of an experiment to determine the best way to do this, about 500 photographs were chosen from the period 1947–48, when Mountbatten was Viceroy of India, and transferred to videodisc using Sony analogue WORM technology. Many of the text documents from this period have already been made available in electronic form.

In this way a multi-media system has been created to provide the user with a large database of text, graphics (maps, scanned images etc.) and video documents, which can be accessed directly or through the hypermedia network of links provided through Microcosm. (The creator of the links might be an archivist, a historical researcher, a teacher or a student. The nature of the links available to the reader will be determined by a set of filters within Microcosm.) The directed paths through the information and tutorial element of the system is provided by HiDES (Historical Document Expert System) software. The HiDES has been developed within the History Department at Southampton to enable students to explore source material bearing on their subject with tutorial support. It has been in use for several years as a teaching tool and in its initial form runs under DOS on a IBM PC-AT or PS/2 machines. In a HiDES package, the student is first presented with the source materials, and then examines them by means of various techniques. For example, the tutor provides a dialogue to stimulate critical discussion by devising questions meant to act as pointers to the debate and to influence its course, and formulates a range of possible replies which might be expected from the student, together with the tutor's own response to each possible reply.

A collaboration has been established between the Microcosm and HiDES teams to explore the three-level model for the use of multi-media information systems in education and research described above. The HiDES software is run in a window under the Microcosm control program, which allows dynamic data exchange between Windows processes. The system is set up so that the student is initially presented with a sequence of documents that are deemed by the historian-author to be crucial to the Indian National Army crisis of 1947. A full account of the historical basis for this application is given in Colson & Hall forthcoming.

Most of the documents are taken from the archive and could be accessed directly through the database or via links, but by making them available as HiDES document files the author is bringing them to the attention of the student as being vital to any consideration of the topic under discussion. At this point the student can also directly access the HiDES question file which provides a preamble stating the argument of the historian in question, posing alternative interpretations and the question to be answered by the student. A file of 'resources' is also available to the student, which suggests references (accessible via links in Microcosm) to information in the archive that the student should also consider in answering the question. The student is free to range through all the material in the database for further evidence. They can use direct database access to any document or they can follow the generic or specific links that have been set-up within Microcosm.

Currently the students write their answers in a 'notepad' (a simple word-processor is used to create text files), which is then analysed by the HiDES software, but in the future it is planned to assess the students knowledge and analysis of the problem by analysing the route they have taken through the information system as well as their written answers. Students' answers will also involve the creation of links that show the chain of argument that they have followed to reach their conclusion. This is an example of the three-level model being applied to a historical application, but clearly the model described above could be extended to any subject area. We believe this will provide an extremely powerful and motivating learning environment, and may point the way to the successful integration of multi-media information systems in education and research.

Conclusions

In this chapter we have presented a case for the use of multi-media technology in publishing and education and have proposed a three-level model for the design and implementation of software environments for multi-media information systems. The three levels are the database, the hypermedia network and the directed path and tutorial applications. Readers are free to interact with the information in the database, or follow links set-up by authors or other readers, or to follow a guided route through the information which may include questions and problem solving activities to help them assimilate unfamiliar concepts and interrelationships. This model assumes a resource-based approach to the provision of multi-media databases for education. The same information can be used in many different ways and for many applications, which helps make the creation of such resources more cost-effective. It is also argued that such an environment will be stimulating and motivating for both students and researchers.

Archaeology is a subject that will greatly benefit from the availability of multimedia information systems. So much of the subject is visual in nature, and cross-referencing between data of different types—text documents, site reports, maps, photographs, statistical data, and more recently video material—is an integral part of an archaeologist's work. The same information systems can be used in schools and colleges to educate and train the archaeologists of the future, as well as forming the formal publications of data. However, large multi-media systems cannot be created overnight and we must begin work now to create the databases that will form the basis of these systems. This will involve much collaboration between archaeological centres, both nationally and internationally, to avoid duplication of effort and to ensure that the data collected is of the highest possible standards in quality of content and presentation. There are also many problems to be solved to ensure that the data and the software created to access it is truly portable. This means working towards agreed standards for the presentation of data, and ensuring the separation of data and software wherever possible.

Many of the remaining problems with book simulacra or educational material are technical, concerned with issues such as:

- Inadequacy of screen size; the 'page size' is not easy to read, and normal screens are unpleasant to look at for long periods.
- Disk space; at present, even simple books like the *Electric Rough Ground Farm* would occupy so much disk space that they would have to be distributed on a medium such as CD-ROM. Since annotation and changing are vital, this would mean maintaining a shadow book on a writeable medium, increasing the complexity of the software.
- The speed of display for adequate colour pictures is not good enough; this *is* being solved, but today's technology is not good enough.
- Despite strenuous efforts, onscreen formatting is inadequate compared to sophisticated traditional typesetting tools.
- The work involved in processing illustrations is non-trivial. We would hope to have a much higher proportion of digital data from new archaeological projects, but we lack the standard archival formats which are now accepted for text. In addition, we must still develop technology to deal with reports being written now, in which illustrations are drawn by hand.

We can see solutions to these problems. The distinction between a traditional database and continuous technical writing is likely to break down, and a single source for both typeset and hypertext documents is well within our grasp. The experiments with the *Electric Rough Ground Farm,* which have so far only covered some obvious textual facilities, show that the greater challenge is in deciding on what model of learning we are hoping to follow.

There is little doubt that the development of multi-media computing systems will significantly affect all aspects of our work, both in the field and in the classroom, in the 1990s. Now is the time to start collecting, collating and classifying the data to take advantage of the new systems as they emerge.

Acknowledgements

Sebastian Rahtz and Wendy Hall wish to acknowledge the support of Professor David Barron for this area of research at Southampton, and thank their colleagues Andrew Fountain, Les Carr, Ian Heath, Gerard Hutchings, Hugh Davis and Frank Colson for stimulating discussions in the past few years. Important critics of this chapter have been Paul Reilly and Brian Molyneaux. To them we owe many thanks.

Sebastian Rahtz's attendance at the Second World Archaeological Congress, where the basis of this chapter was presented, was made possible by a grant from the British Academy, and by financial assistance from the University of Southampton.

Notes

1 This work was originally presented as two papers at the second World Archaeological Congress, as separate discussions of electronic reports, and of the possibilities of videodisc. It is the result of research work carried out in the University of Southampton, in collaboration with the Oxford Archaeological Unit.

2 The writer's self-published visionary volume, *Dream Machines*—Nelson 1987—is a good introduction to his ideas, and to some of the intellectual challenges in store for us in an electronic world.

3 The *Electric Rough Ground Farm* is unpublished in any formal sense; the demonstration system can be supplied to interested parties on application.

References

Allen, T., T.Darvill, S.Green & M.Jones 1991. *Excavations at Rough Ground Farm, Lechlade, Glos: a prehistoric and roman landscape.* Oxford: Oxford University Committee for Archaeology.

Alvey, B. 1990. Graphic interpretations of stratigraphy using desktop computers. In *Precirculated papers for Information Technology themes at World Archaeological Congress 2, Venezuela, September 1990, Volume 1: data visualisation,* 1–6. Winchester: IBM UK Scientific Centre.

Boast, R. & D.Chapman 1991. SQL and hypertext generation of stratigraphic adjacency matrices. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 43–52. Oxford: British Archaeological Reports.

Brand, S. 1988. *The Media Lab: inventing the future.* Harmondsworth: Penguin Books.

Bush, V. 1945. As we may think. *Atlantic Monthly* **1(176) July,** 101–108.

Colson, F. & W.Hall forthcoming. Multimedia teaching with microcosm-HiDES: viceroy Mountbatten and the partition of India. *Journal of the Association of History and Computing* **3 (2).**

Desachy, B. & F.Djindjian 1991. Matrix processing of stratigraphic graphs: a new method. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 29–38. Oxford: British Archaeological Reports.

Fountain, A.M., W.Hall, I.Heath & H.Davis 1990. Microcosm: an open model for hypermedia with dynamic linking, Technical Report 90–13. Department of Electronics and Computer Science, University of Southampton.

Frere, S. (ed.) 1975. *Principles of Publication in Rescue Archaeology: report of a working party of the Committee for Rescue Archaeology, Ancient Monuments Board for England.* London: Department of the Environment.

Gardin, J.-C., O.Guillaume, P.-Q.Herman, A.Hesnard, M.-S.Lagrange, M.Renaud, & E.Zadora-Rio 1988. *Artificial Intelligence and Expert Systems. Case studies in the knowledge domain of archaeology.* Chichester: Ellis Horwood.

Goldfarb, C.F. 1990. *The SGML handbook.* Oxford: Oxford University Press.

Hall, W. 1988. The use of interactive video in undergraduate teaching. In *Aspects of Educational Technology XXI* , Mathias, H.N.Rushby, & R.Budgett (eds). 158–163. London: Kogan Page.

Hall, W., P.Thorogood, B.Sprunt, L.Carr & G.Hutchings 1990. Is hypermedia an effective tool for education? In *Proceedings of the World Conference on Computers in Education,* McDougal, A.L. & C.Dowling (eds), 1067–1071. Amsterdam: Elsevier.

Herzog, I. & I.Scollar 1991. A new graph theoretic oriented program for Harris Matrix analysis. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 53–60. Oxford: British Archaeological Reports.

Hodder, I. (ed.) 1989. *The Meanings of Things: material culture and symbolic expression.* London: Unwin Hyman.

Huggett, J.W. & M.A.Cooper 1991. The computer representation of space in urban archaeology. In *Computer Applications and Quantitative Methods in Archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 39–42. Oxford: British Archaeological Reports.

Lock, G. & C.Dallas 1990. Compact disc-interactive: a new technology for archaeology? *Science and Archaeology* **32,** 5–14.

Martlew, R. 1992. The implications of large scale image storage for primary archaeological research. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 21. London: Routledge.

Maytom, J. & K.Torevell 1991. Putting the public in the picture. In *Computer applications and quantitative methods in archaeology 1990,* Lockyear, K. & S.Rahtz (eds), 21–22. Oxford: British Archaeological Reports.

Nelson, T.H. 1974. *Dream Machines: new freedoms through computer screens—a minority report/ Computer Lib: you can and must understand computers now.* Chicago, Illinois: Hugo’s Book Service.

Nielsen, J. 1990. *Hypertext and Hypermedia.* London: Academic Press.

Rahtz, P. & S.Hirst 1976. *Bordesley Abbey, Redditch, Hereford-Worcestershire: first report on excavations 1969–1973.* Oxford: British Archaeological Reports.

Rahtz, S., L.Carr & W.Hall 1989. New designs for archaeological reports. *Science and Archaeology* **31,** 20–34.

Rahtz, S., L.Carr & W.Hall 1990. Creating Multimedia Documents: hypertext-processing. In *Hypertext: state of the Art,* Green, C. & R.McAleese (eds), 183–192. Oxford: Intellect.

Ryan, N.S. 1988. Browsing through the stratigraphic record. In *Computer and Quantitative Methods in Archaeology 1988,* Rahtz, S.P.Q. (ed.), 327–334. Oxford: British Archaeological Reports.

Schneiderman, B. & G.Kearsley 1989. *Hypertext Hands-On: an introduction to a new way of organising and accessing information.* New York: Addison-Wesley.

Stutt, A. & S.Shennan 1992. Designing a workbench for archaeological argument. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 18. London: Routledge.

Wilcock, J. 1992. On the importance of high-level communication formats in world archaeology. In *Archaeology and the Information Age,* Reilly, P. & S.Rahtz (eds), chapter 8. London: Routledge.

Yankelovich, N., B.Haan, N.Meyrowitz & S.Drucker 1988. Intermedia: the concept and construction of a seamless information environment. *IEEE Computer* **January,** 81–96.

Index