Comp. Appl. Arch. 13, 1985.

ARCHAEOLOGY, DATABASES AND MICROCOMPUTERS

Jonathan Moffett

Institute of Archaeology, 31-34 Gordon Square, London WC1H 0PY.

Introduction

This paper describes the conclusions of a three year investigation into the management of archaeological data on a microcomputer. The purpose of the research was neither to develop a Data Management System (DMS) from scratch, nor to compare different existing commercial DMS packages. This research concentrated on the use of one computer and one commercial package and applied it to three archaeological situations: an excavation, a Sites and Monuments Record (SMR) and the management of radiocarbon dates.

The Computer System

The computer system used for this research was a Sirton Computer Systems Midas 3HD, which is a Z80A microprocessor-based microcomputer. As such it was able to run the CP/M 2.2 operating system. Table 1 gives other details about the machine. The DMS package used was MDBS I (MDBS 1980), a network Database Management System (DBMS). Table 2 gives details about this package.

Table 1: Computer System

```
        4 MHz Z80A microprocessor
        64 Kbytes of RAM
1 x 20 Mbyte Winchester Hard Disk
1 x 1 Mbyte Floppy Disk
2 x RS232C serial interfaces (apart from console)
```

Table 2: MDBS I: A Network Database Management System

```
Maximum No. of Record Types         = 254
Maximum No. of Fields per Record    = 255
Maximum Length of a Field           = limited by page size
Maximum No. of Sets                 = no limit
Maximum No. of Records per file     = limited by disk storage
Maximum No. of Usernames/passwords = 255
Security in form of usernames and passwords and 255 levels
of read/write permissions.
```

The decision to use these was taken in 1981 when this hardware and software were considered to be the most suitable then available. Thus, the Midas 3HD was chosen because it was compatible with the machine used by the post-excavation organisation which supplied data for the first case study, it used CP/M, and also it had a 20Mbyte Hard disk.

MDBS I was chosen over other DMS such as dBASE II (Ashton-Tate 1980) as these were relational-type DBMS with several limitations regarding the number

of fields in a record or the number of characters in a record. MDBS I was a more flexible system because the constraints tended to be imposed by the computer system. An example of the difference is that dBASE II files are allowed up to 32 fields. MDBS I in theory can have 254 x 255 fields. In practice the allowed number is much less. The choice of a network DBMS also ensured that a database structure could be designed to fit the data, rather than the data having to fit a specific structure. However, as it turned out, all three case studies ended up with the same type of network structure (Moffett 1985).

Case Study I: An Excavation Database

Catton, Jones & Moffett (1981) described the initial phases in the development of the Mucking Excavation Database. Mucking, a multi-period site, proved to be a never-ending source of data and although this was initially considered to be useful by the end of the research period there was too much data for the computer system to handle. Furthermore, data were still being entered on the computer. At the end of the investigation only a small part of the data had actually been loaded into a database. These were the cross-references between notebooks descriptions and coordinates. This file alone covered some 4Mbytes and constituted some 26,000 unique coordinates and 13,000 unique pages of references. Additional coordinates and references were anticapated due to the cross-references in the find and feature data files. Table 3 gives details of the quantities of the range of data involved.

Table 3: Examples of the data produced by the Mucking excavations.

Data Type	No. of Bytes	No. of Fields	No. of Records
Animal Bone	81	5213	657408
Anglo-Saxon Burials	75	781	123008
Anglo-Saxon Pottery	80	1989	585600
Charcoal	13	6333	392320
Fired Clay Artefacts	81	500	70016
Fired Clay Blocks	38	45	5248
Fired Clay Scraps	167	11172	700416
Flint	16	6421	553600
Iron	43	779	97152
Notebooks	10	19316	1654912
Notebook Catalogues	16	1078	317568
Prehistoric Pottery	56	(to be entered)	
Plans	14	3972	1028096
Romano-British Burials	75	157	27008
Romano-British Pottery	97	15535	2230655
Round Houses	21	106	113920
Slag	21	1957	134784
Soil and Raw Clay Samples	22	1540	115200
Stone	38	1114	72704
Textiles	12	145	44288
Tile	40	4233	445696
Trenches	44	60	49920
Totals =	1060	82446	9419519

The purpose of the DBMS was basically to provide a complete cross-reference catalogue of all the finds and features across the site. However, the

archaeologists had developed no strategy to determine what they wanted the computer to do for them. Although results such as distribution plots and pie charts have been produced from the Mucking post-excavation data, the application of the DBMS failed for two fundamental reasons:

No pre-planning of what the computer was to do
The quantity of data was to great for the computer

The second is probably a consequence of the first in that the use of the computer needs to be well thought-out before one is actually bought.

Case Study II: An SMR Database

The author was given the opportunity to develop a Sites and Monuments Record Database for the county of Bedfordshire (Moffett 1984). The Conservation Department of the County Council bought a Midas 3HD which was identical to the one used by the author so that the DBMS developed on one system could be automatically transferred to the other.

In contrast to the excavation database above, this application can be considered successful, basically for the opposite reasons that the previous database was not. The requirements of a computer-based system had been well thought out prior to the decision to buy a computer, while the amount of data was relatively small, at least compared with the excavation. There were initially some 13,000 unique primary records each of which had a maximum of 13 fields, some of which could be repeated. MDBS I allowed the development of a general network structure which could efficiently accommodate this data (Moffett 1984).

The purpose of the DBMS was to allow for the retrieval of selected primary records and also for their continued updating. The original SMR was based on a record card to store the basic information relating to a particular site, monument or building, a longer descriptive record card describing the site in greater detail, and an optical coincidence card system, for searches. The computers immediate task was to replace the optical coincidence cards, but in the long run it may be possible to store the entire record on a computer, as descriptive digest files are to be stored on the computer, one for each primary record, while photograph and slide references are also to be stored. With a projected maximum of some 25,000 primary records, and a minimum size of 2k per digest file, this results in an eventual need for 50Mbytes of storage space. In terms of floppy disks, and including second back-up copies, at least 100 x 1Mbyte floppy disks will be needed. The database file itself is projected to reach about 13Mbytes in size, while ancillary files may take up another 3 or 4Mbytes. Thus for a relatively simple archaeological situation the required computing capacity can easily become greater than the current capabilities of microcomputers.

Case Study III: a radiocarbon database

The radiocarbon database has been described elsewhere (Moffett and Webb 1983), although that design has now been replaced by a structure similar to that of the SMR database (Moffett 1985). There are some 2400 Palaeolithic and Mesolithic dates for the Old World, which were initially transferred from a card index sorted by site name. The principal purpose of a computer record was to make additions and corrections to the date list easier. Also it allowed for quicker and more thorough searches, with the option of producing graphical displays of the dates.

The small number of dates, this database is about a fifth the size of the SMR, and the small number of fields (16) suggests that MDBS was like using a hammer to crack a nut and that a dBASE II-like solution would have been as suitable. The dates have now been transferred to a multi-user mini-computer at the Institute of Archaeology and are stored using Informix a relational DBMS.

General conclusions

Three conclusions can be drawn from this research:

The hardware used during this research was reliable but it was not an easily transportable machine. However, portable machines such as the Epson PX-8 are not designed to be used for such tasks. The Kaypro 10, which has a 10Mbyte Hard disk may be useful for small scale databases. The use of CP/M allowed access to a considerable range of software, although the limitation of 64K of memory is not sufficient for large database tasks. It is also very easy to use all 20Mbytes of disk space, when generating large databases.

The software MDBS can be applied to archaeological situations, but it is too complicated to be implemented by most archaeologists who do not have time for the intricacies of databases. In this respect dBASE II or a similar relational-like DBMS which are menu-driven would be suitable for archaeological work. MDBS I is most suitable for medium-scale database problems, which would be inefficiently solved by the relational method, smaller database with less than 5000 records can probably be satisfactorily handled by dBASE II. MDBS I is now no longer available, although MDBS III and Knowledgeman are produced by the same company.

Databases in general are a necessity for Archaeology. Although they tend to be static in structure they do increase in size. Updating of records is not as common as retrieval operations and so archaeological databases should be designed with this in mind. Having said this, one of the two primary functions of the SMR database was the updating and addition of data, and so this function cannot be ignored.

The future

The nature of archaeological computing has changed considerably over the past four years. Not only is there a greater range of computers as well as software, the power of these computers is greater than that of their predecessors. Hard disks have now become common-place, and while the 8-bit still has a place in computing, the 16-bit microcomputers are slowly superceeding these. Furthermore, the arrival of such computers as the Amstrad 8256 and the Amstrad 6128, for under £ 500 opens up a whole new computing world for low budget computing.

References

ASHTON-TATE. 1981 dBASE II, Users' Manual. Ashton-Tate Ltd., Culver City, California.

CATTON, J.P.J., JONES, M.U. & Moffett, J.C. 1982 The Mucking Excavation (1965-78) Database. Computer Applications in Archaeology 13: 36-50.

M.D.B.S. 1980 MDBS I, Users' Manual. Micro Database Systems Inc., Lafayette, Indiana.

MOFFETT, J.C. 1984 The Bedfordshire County Sites and Monuments Record Database Management System. Computer Applications in Archaeology 12: 101-9.

MOFFETT, J.C. 1985 The Management of Archaeological Data on Microcomputer. Unpublished PhD dissertation, University of London.

MOFFETT, J.C. & WEBB, R.E. 1983 Database management of radiocarbon dates. In Aspinall, A. & Warren, S.E. (eds) Proc. XXII Int. Symp. on Archeometry & Arch. Prospection, Bradford, 67-72.