# THE FIFE ARCHAEOLOGICAL INDEX - A COMPUTER IMPLEMENTATION

J.B. Kenworthy Lecturer in Archaeology
J.R. Stapleton Computing Laboratory

J.H. Thurston Information Systems Unit

University of St. Andrews,
St. Andrews
Fife.

## 1. Introduction

This paper describes a database, the ISl Filing System, developed by the Information Systems Unit at the University of St. Andrews, and its application to the formation of a sites and monuments record for the Fife Region. Several problems associated with the use of computerised databases by archaeologists are discussed.

## 2. The Fife Archaeological Index

In common with the rest of Britain, destruction of archaeological sites in Fife is increasing from year to year. Only by the provision of adequate information to planning authorities, developers, and local archaeologists, can this problem be alleviated. Ideally, the record should be built up by a Regional Archaeologist working within the Regional Planning Department, but in the current financial and political climate this has so far been impossible. The work has therefore been undertaken within the University of St. Andrews with financial help from the Russell Trust and the Carnegie Trust for the Universities of Scotland. When complete, the index will consist of a set of marked Ordnance Survey 1:10000 maps linked with a cross-referenced card-index, holdings of relevant plans and air and ground photographs. For ease of handling, the information will also be held on the computer, as described below. Compilation has started, but it will be several years before all the information will have been checked in the library and in the field. As an interim measure, the information contained in the OS Archaeology Record Cards is being used to provide a point of departure, to be supplemented as work progresses, with

the addition of information of possible threats.

The aim of the index is threefold:

(1) To allow archaeology to be integrated into the planning process by means of close co-operation with the Planning Department. It is hoped that eventually a copy of the index will be held there;

(2) To provide background for field surveys in areas of immediate threat, and by analysis of data, to identify areas where particular archaeological problems arise, and where intensive field-work is necessary;

(3) To provide a basic source of information for all those carrying out research work in the area.

Why use the Computer?

Doubts have been expressed as to the usefulness of computer-based information retrieval systems, especially as regards their time-effectiveness for small jobs. It is intended that the card-index will be used for small requests, where the time taken to retrieve the information from the computer would have allowed the retrieval of several items by hand. But this does not mean that the computer is useless outside the Planning Office. Although in our case, the amount of data is relatively small, even with the inclusion of artefact finds it was felt from an early stage that the number of variables concerned, and the types of questions to be asked of the data, necessitated a computer-based information retrieval system.

Our requirements were roughly the same as those propounded by Cutbill (1974), though the manual operation of small jobs means that terminal use is not of the greatest importance. On the other hand, the most important considerations were:

(1) The ability to handle sizable amounts of text - if the computer print-out contains all the information necessary, much copy/typists time is eliminated and errors are less likely to occur. This is where the computer can be of great use, as also in the production of distribution maps for any combination of variables - a very time-consuming process by hand.

(2) Simplicity of data input, since the system is to be operated by people with relatively little computer-experience.

(3) Simplicity of 'coding'. The output should be immediately comprehensible to those not conversant with the system, without the use of a cumbersome 'code-book'. This should also minimize error.

Used for sizeable jobs, the computer-based system should repay amply the time spent in initial data-entry and programming. The research applications of the system should also be apparent.

An approach to the University Computing Laboratory showed that the database described below, already in use for University records, would be suitable, despite certain restrictions, which are negligible compared to its advantages. The implementation of the system for the Index is described below.

3. Implementation

An archaeological site may be described by considering a series of attributes. To show how this is done, consider the following site:

| | | |
|---|---|---|
| Site Name | : | Wester Nether Urquhart |
| Site Type | : | Cairn |
| O.S. Card No. | : | NO 10 NE 9 |
| N.G.R. | : | NO 185 086 |
| Height O.D. | : | 400 ft. |
| Period | : | Bronze Age |
| Major Attributes | : | None |
| Associated Finds | : | None |
| Condition | : | Destroyed |
| Size | : | Not Known |

This is not a full description: obviously Parish, County, Land Use, etc., could be added. It may be that the above was all the information available from the source document. A search would then be made to see if further information is available. The result of the search could provide information on the Parish, County, Land use, Land owner, aspect, slope of site and bibliographic references. With the proposed database the addition of this further information under attributes not yet used for the given site poses no problem. In fact, the record may be extended up to a maximum total length of 10,000 characters; thus if information is available which does not fit into any particular attribute, a 'Description' attribute may be used. As a further benefit of the database system, if at any time it is realised that another attribute, not yet defined, is required, this new attribute may be added to the system without any need for a complex re-generation of the file. At present some 60 attributes are in use with the system.

In the attributes listed in the above example, it will be noticed that they are very similar to those used in a feature or finds index on an excavation (for example, Site type - Feature type, N.G.R. - Site grid reference, Height O.D. - Height from site datum, Period - Level, etc.) In this way an index for an excavation may be generated and used for keeping track of the progress of the excavation and its finds. Such a database would probably be kept separate from a county index while the excavation was in progress; it could be merged into the county index when

the excavation had finished.

As noted below, the database system itself is limited in its statistical capabilities to the production of counts of requested classes of objects. However, by the use of external linkage it would be possible to add a statistical package to the system. For simple operation it would be possible to save one of the intermediate lists of the requested records produced by the system and to analyse this to produce, for example, a distribution map.

Since the attributes are entered in 'English' it will be appreciated that simple typing errors are not fatal and can easily be corrected (see database description). However, if a great deal of similar data is to be entered, e.g. a large number of similar artifacts, there is no reason why the data should not be coded in those areas where it is repetitive. The data can then be decoded by use of the automatic data correction program (see 4, iv (c)). This process would reduce entry time; care must be taken, however, to avoid the dangers of transposition errors in numerically coded data.

## 4. The ISl Information System

### i) Introduction

The ISl Information System is a general purpose filing system developed by the Information Systems Unit at St. Andrews University. It is written in a machine independent high level language - Snobol 4 - and is in principle easily transported to another machine environment. In fact, the batch operating system through which the system interacts - 44MFT - is the main restriction on its use by other computing establishments.

ISl began its life as an experimental data base; the principles, ideas and experience gained through it are in no sense restricted and the designers will be very pleased to discuss the database, either proximally or distally, with interested groups or individuals.

### ii) Working Environment

a) Machine configuration: ISl demands are minimal. Nevertheless, it is unlikely that the system could be run on a mini machine; the main restrictions are that there be enough backing store to hold up to six

work data sets and enough store to handle a full sorting package. The system has been developed on an IBM 360/44 with 256K. Normally 112K of the store is required for running the system together with two dedicated exchangeable 7M byte discs. Tapes are desirable though not mandatory. Output is normally on a line printer.

Regardless of the central and peripheral computing facilities available, ISl accepts any reasonable method of data entry, viz. paper tape, punched cards, key to disc, key to cassette, key to tape. Whatever medium is used necessitates the same procedures and standard syntax for data.

b) Personnel: To run ISl successfully at least one competent programmer is required. All data entry and data retrieval, on the other hand, may be handled by non-specialist staff. We have found that any competent typist becomes fully conversant with data entry and data correction procedures following one hour of familiarisation.

iii) File Structure

The file consists of an ordered set of records. Each record describes one item, the record consisting of an unordered set of attributes which together fully describe the item. In the case of the Fife Archaeological Index, an item will be a find; that is, each record pertains to a different find. The attributes comprising the record will be facts or data relevant to the find.

Each attribute actually consists of two parts: the attribute code and the attribute value. The code is a tag defining the relation of the value to the item. It is necessary since the attributes are unordered and all attributes need not be present in any given record. The value is a string of characters representing the value which this item takes for the specified attribute. Both the attribute code and the attribute value are terminated by delimiters; either can thus be of variable length.

The diagram below shows schematically how a part of the Fife Archaeological Index would be held under ISl. The Attribute column is for reference only and is not held within the records.

| Attribute | Code | Value |
|---|---|---|
| Reference Number | RN | F136136 |
| RCAHM Number | RCN | 513 |
| O.S. Card Number | OSN | NO 10 NE 9 |
| N.G.R. | NGR | NO 185 086 |
| Parish | PH | Strathmiglo |
| Site Name | SN | Wester Nether Urquhart |
| Site Type | ST | Cairn |
| Period | PE | Bronze Age |
| Major Attributes | MA | None |
| Associated Finds | AF | None |
| Size | SZ | Not known |
| Height | HOD | 400 feet |
| Aspect | AS | North Slope |
| Land Use | LU | Under Cultivation |
| Condition | CN | Destroyed |
| Bibliography | BIB | Small, A (1823) Roman Antiquities in Fife (1823), 217 |

iv) System Programs

The functions of the programs which comprise ISl are:

Updating the data

Data retrieval

Automatic correction of data

Linking to other software, e.g. external statistical packages.

a) Updating: This process takes two forms: firstly, amendments to attribute values within existing records on the file; and secondly, addition and deletion of records and the addition, deletion and alteration of attribute codes. The latter process is normally only performed by the person responsible for the running of the system. The former process is the main task of data entry personnel. Data is entered in an approximation to ordinary English. No coding of data is required although it is not forbidden. For each 'fact' or attribute value to entered or amended, an attribute code must first be given. The codes are often initials e.g. SC - Site Condition. The value would be entered following the code. As discussed above, any medium may be used for the recording of amendments - punched cards, paper tape, etc.

The Updating program coupled with the appropriate preprocessor will

incorporate the amendments into the file.

b) Retrieval of Data: The second function is performed on ISl almost entirely by a single program package. By the skillful use of a set of driver statements, any conceivable listing of data, presented in any conceivable order, in any conceivable format, is available within a few minutes; subject that is, to the obvious constraint that the data is extant on the database.

An example of a set of driver statements is given below:

```
TEST  PE = '15th century'
SORT  (SZ,20)
TITLE '15th Century Erections in Fife'
LIST  (40,SN,SITE NAME)  (20,NGR,GRID REFERENCE)  (20,SZ,SIZE)
```

This will give a listing in height order of all 15th century erections in Fife, listing in addition OS grid references and height.

c) Automatic correction of data: This facility is invaluable where non-specialist data entry staff are used. Its main uses are for correcting typing errors or expanding data coded to facilitate data entry. A single statement can alter any or all the occurrences of any value on the file.

d) External linking: The system allows for unlimited use of external packages, programs, etc. Typically, a high powered statistical package may be linked into the system to provide specialised facilities not present in the database. The programmer would be responsible for the linking programming. Special purpose preprocessors or listing packages could also be used in the same say.

Clearly, a summary as brief as this can only superficially describe an information retrieval system that has taken three man years to design and implement. Because also of the somewhat esoteric nature of the language in which it is written, this review of the system must perforce be in usam Delphini.

Reference

Cutbill, J. (1974) Computer-based filing systems. Computer Applications in Archaeology (1974), 81