DISTRIBUTED PROCESSING FOR KARST

G. Pilkington

Most record–keeping began as individual diaries, progressed to club journals with trip reports, and stopped there. Australia is one of the few regions that has developed a collective, systematic records system that crosses club boundaries.

Like most large scale integrations, the standardisation of ideas progresses slowly. We now have metric scales and measurements, common map symbols, a systematic cave identification scheme and a common karst nomenclature as represented by the Australian Karst Handbook (AKH).

Have we reached the limit of co-operation? We have not. What we need now is a better, faster, coherent method of collating and presenting our data to facilitate data exchange, such as for the AKH, as well as making cave data searches easy.

Historically, speleological society records have been designed or defaulted to manual methods. Present Australian data storage systems are manually oriented, even when stored in computer readable form as the AKH will be. Survey reduction programs exist for computers, but their data banks tend to be designed as cards and independent of all other cave data. What computer data handling exists has been developed in a very haphazard way by individual enthusiasts who use computers regularly. Each program tends to be written and run independently, but might depend on serial execution and hands-on formats. The only developed branch of data manipulation that I am aware of is cave survey reduction and plotting.

The real problem is NOT that computers are available only to the privileged few, because desktop versions with random access file capability are now readily available. It is that they are still inconvenient to use and take a long time to program with quality bug-free data handling systems.

A smaller but still significant problem is the task of transferring even new data to computer storage. With the manual system you only have to slip the relevant piece of paper or its photocopy into the right cave folder but with computers it is often necessary to retype each article. While the original documents are still handwritten or typed we cannot afford the expense or time of imitating the manual system. Word processors are now reaching down to our level of use. If people start compiling their newsletters and journals etc. on these devices, then the problem of data transcription transforms into one of retrieval and indexing. For now let us concern ourselves with the smaller task of storing just summary data — an expanded form of the AKH Cave Summary Sheet data.

Why do we need computer storage for our data? It is the responsibility of the societies to supply data to the Australian Karst Handbook on forms suitable for computer input. These take a long time to fill out. It would be to each society's advantage if that effort was used to create a data file for their own use that could generate the AKH data on request. It would enable rapid updates of the AKH as well as being available for answering local questions which are normally only answerable by the most knowledgeable local cavers or after a week–long search through club records.

Why do we collect data in the first place? It is mostly out of interest in the data itself or in some visual equivalent like a map. This simplifies the requirements of a computer

Caveconvict — Proceedings 13th A.S.F. Conference, 1980

storage system because both input and output are similar in form and, therefore, the same type of storage is good for both updating and for producing required answers.
The data storage system should be able to accept the following types of data:
Cave summary
Map summary
References
History
Survey

It is not necessary to have one record type as long as adequate cross–indexing is used; this applies equally well to manual files.
While storing and retrieving the data can be done on almost any machine with any kind of input/output, the best system is an interactive one that can verify data by direct comparison to the data file and ask for the correct values at the time of entry. Batch input is possible but frustrating when a small error prevents further data processing. An ideal system would have:
interactive on-line access to the data bank
a random access data bank
a tolerant "humanised" special purpose language (SPL)
visual display unit and keyboard
printer
digitizer and plotter

Not all these would be necessary for any one society; a subset could be used with alternatives such as card input. Special facilities such as word processors for writing and editing would enable automatic keyword and reference extraction to be included in the data bank.
If we use an SPL for interaction with the data bank, the difficulty of using computers can be minimised. A uniform common language could be written that would operate for all machines and hence be identical for all societies. This would make data exchange and information requests between societies very easy. Examples of possible commands are:

```
PLOT 501 AT 1:200 ON A4
STORE UNDER 6N37:  Bats seen 2/1/66; White Lake dived 1980
REFERENCE 6N37:  HILL AL, 1966.  Mullamullang Cave
expeditions 1966.  CEGSA Occasional Paper 4.
```

Present programs could be spliced on by adapting the SPL to call them — generating the correct input for the particular programme.
The execution of each command can be written in structured language which is also machine independent. For example:

```
IF COMMAND starts with PLOT then execute PLOT routine
ELSE IF COMMAND starts with STORE then
etc.
```

The final writing for each machine will then be able to be written from a common source. This means that more effort can be spent in making the structured version "bug-free" and after implementation, errors detected by any user can be easily corrected in the common language and hence in every program written from it. If coded versions of the structured language are written in BASIC and PASCAL, most machines that will be used can obtain ready–made packages, even if they need slight adjustment for each installation.

THE NEED FOR SUCH A COMMON SYSTEM IS NOW, before too much effort has been spent on individual methods, and before the AKH computer version stimulates many unique incompatible data banks.

I therefore advocate that the ASF start an Ad Hoc Commission on Karst Computer Data Banks to do the following:

(A) Decide the best method to achieve a common operating language or SPL.

(B) Design an SPL.

(C) Design a structured language solution to the SPL (an SLS).

(D) Obtain one or more High Level Language (HLL) versions of the SLS.

(E) Control the correction and distribution of the SPL, SLS, and HLL's.

(F) Recommend standard implementation procedures and aid in adapting the system to each society's needs and equipment.

(G) Co-ordinate extensions and revisions to the SPL, including output presentations especially for data exchange.