PROGRESS WITH LERNIE

Leo Biek.

Ancient Monuments Laboratory.
Department of the Environment.

In 1972 I tried to convey to the Oxford Conference on Archaeometry my impression of a certain random flavour about work in that field: What precisely were we doing? How precisely did we need to do it? We were all together spending no more in a year than ERNIE gave away every month - two or three miles-worth of motorway. It would be helpful, even so, if we could recognise some sort of pattern in the work which seemed, rather, to be actuated by the astrological conjunction of equipment, personalities and material! We lacked an automatic device for running our problems against a checklist of those important questions.

In particular, computer techniques were being used mostly as more or less straightforward statistical packages, run in fits and starts on very raw descriptive and experimental data with paper tape and standard print-out. Above all there was no processing of the results for feed-back into planning the next project. We had not begun to think of putting the computer onto the most vital part of the loop: our own ability to think and, especially, to learn. As Doran says below, we are not very good at conveying the large amount of everyday background data which we unconsciously assume all the time.

I went on to suggest that what we needed most urgently was to develop interactive graphics - applied to something like visual typological comparison of artefacts. Basically I wanted to put the radiograph of a Saxon iron knifeblade in at one end and get a hardcopy of outlines and references for all parallels out at the other. But I also wanted to see all these outlines on the screen if I wanted to, and make decisions about classification during run. And I wanted any such decisions as I chose to be absorbed into the program at the same time as the actual new knifeblade data were added to the bank. Audience reaction, expressed privately later, made me go on in odd moments outside real time. I found there was enough suitable software in my head. The original print-out went something like this:-

LERNIE - a Learning Experiential Research Needs Integrating Evaluator for Archaeo-scientific investigation*

(A) Basic feature - Complete rationalisation of all processes based on precedent/comparative analysis/collation/ integration/sorting & sifting/sequential display.....;

to give greater speed in providing basic framework as well as greater security (avoiding reliance on memories and ad hoc checklists) and capacity;

by progressively stepped-up homing programs of trending significant elimination of cumulative data;

involving accelerating multivariate refinement of pattern recognition.

(B) Basic Key pivot - continuous and constantly focussing graphics interaction

(C) Applications - (a) Immediate:

1. Pattern-recognition in analysis of computer recorded and filtered geo-surveys (in progress)

2. Pre-planning of dig orientation....

7. Selection of resultant needs (user study)

8. Experiential analysis (data efficiency, comparative cost/benefit)

Key pivot - Translation and presentation of character pattern

(b) First remove: Development of (a) 8 -

Auto-experiential search for optima/expansion/ refinement.....

by continuous homing minitor evaluation of all cumulative data as they are processed

(c) Second remove: Perfection of factors accelerating (b) -

Increase in homing agility.....

by improvement in methods of significant approximation/ assessing economic determinants/indicating professional trends/refining selectivity*

LERNIE was launched at the 1973 Archaeometry Conference far too early one morning. The very same afternoon Scollar described an experimental run on Grafix 1, designed specially for interactive image processing and pattern recognition. He was interested mainly in rapid and direct visual enhancement of interesting features in magnetic prospecting data. But for me this was the physical nucleus of a basic system on which we could start actually to build up LERNIE.

We now need a breakthrough to effect a change of heart. Into our cosy hunch-and-crunch research routine we need to insinuate operational research on Research, work study on research thinking, feasibility studies on research projects. The ulitmate aim is a continuous radar-like sweep of the whole field of activity - auto-collating all needs, facilities and projects; balancing precision against significance, and evolving real needs linked to actual methods.

This will generate a force field of foresight in which demand and supply will be in dynamic equilibrium. And the system needs to be both highly interactive and self-adaptive, so that we can change our minds even as we are making them up. The next thing is to get a move on. It is not so much a change in speed as in dimension that is required in our thinking: not a change in velocity, but a capacity for constantly changing the rate of change in velocity. We can see it in action on any football pitch.

What we need is to develop a kind of simultaneous-lightning-chess mentality. The only way is to strike up some form of symbiosis with the computer. From the beginning our minds have been resting securely on one basic axiom: One thing at a time. This is now fading fast. If we don't quickly diversify into the unthinkable we shall be drowned in data and alternatives.

We need to think big but we can only start small. My own efforts have been concentrated into three main channels:

1. Collative report writing

Often several scientific specialists will need to examine a single archaeological find, group or feature. The final report will then additionally require creative integration. This can be immensely exciting in its forensic discovery but much of it is boring, tiring and essentially mechanical - an obvious subject for simple computer application.

For a very well preserved head of Roman hair, recently found at Dorchester in Dorset, I had the following specialist reports on different aspects of the evidence:

A. The lead coffin
B. The gypsum filling
C. The textile impressions in the gypsum
D. The human bones
E. The hair - general: morphology, chemistry, preservation
F. The hair - amino acid content
G. The hair - u.v. fluorescence
H. The hair - keratinophilic soil fungi

In addition there were some general background notes, as well as the archaeological account and references to previous finds of a similar nature.

Briefly, a primary card was made out for each statement in its essential form, viz., "Impression in gypsum of fine plain weave textile (size & count, etc.) with folds." This message was written on a conventional edge-punched card which was then clipped under three main headings:

Material - here, "gypsum" and "fibres";

Type of data - i.e., "present evidence" - as against "present inference" or "others' evidence";

Key concept - size and numerical data, cultural, chemical, technological, preservation, etc.

There were also the usual codes for back reference to source, as well as some refined criteria such as levels of confidence or precision, and of appraisal ("primary", =visual; as against "secondary", =analytical; or more complex investigations).

These primary cards were needled out first by material within one report, then by material across reports, and finally by various key concepts in turn.

Such a selective mini-matrix analysis ensures that all 'connections' are presented for appraisal of significance. This in turn reveals both 'new' and 'more jomplex' connections which are similarly analysed. Thus the fineness of the textile impression emerged as a highly informative factor with wide repercussions - not least in regard to the manner in which the gypsum entered the coffin. And the lead oxide detected in the head stain on the coffin bottom by one specialist is set, however fleetingly, aginst lead oxide in a 'natural' hair dye quoted by another.

At any chosen stage the report can then be left to write itself or can be further elaborated. All the cards remain source material in a classified form, including automatically all the associations made during the exercise itself, and the whole process remains flexible and cumulative within the size limits of the card. It has also been used in a variety of other contexts, from analysis of data on gypsum burials and on preserved hair, to the ordering of day-to-day priorities in the light of constantly changing states of 'urgency' and 'importance'. The continuing 'program modification' in selecting priorities has deeper implications, for instance in defining ever more precisely the relative and variable importance (to self) of 'other people'.

The whole system can suitably translate to some simple package such a IBM's MINIMIS and then becomes a totally automatic 'first draft' producer. But for the moment it may be wise, in general, to continue on cards; there are, for example, 'hidden' associations which are (literally) written into a card, and received back on seeing it - both automatically - but which cannot yet be programmed efficiently. This was very plain to me in using the system to prepare this report - i.e., what you are reading this moment - and to relate it as far as possible to the contributions above and below.

2. Artefact outline comparison ("Parallelography").

The simplest way of doing what I suggested at the beginning would seem to involve a scanning micro-densitometer head linked to the computer through a suitable discriminator. The stepping motors would automatically provide the digital information (x,y) for the matrix. The logic input would depend on the type of discrimination. A low level discriminator could simply 'reproduce' the x-radiograph (light=1, dark=0). But far superior powers of interactive comparison would be provided by some logic linked to a change in film density (crossing light/dark = 1 = crossing dark/light). This would produce an 'empty' outline into which others could be fitted. This has obvious advantages especially when the objects differ in size and the smaller can be enlarged inside the larger in proportional matching.

AEP International Ltd., have recently developed a "Scratch Pad Memory" which does much the same thing with an SEM, by continuously refreshing the elemental destribution image which appears on the CRT, and which thus can actually be observed to build up on the screen. The module can also concentrate on selected areas and provide image enhancement.

"Parallelographs' could deal with radiographs of any artefacts with characteristic outlines ; the most obvious types of object, apart from various cutting implements, include buckles, horseshoes and keys. For well-preserved objects in other materials a suitable, 'contrasty' black-and-white negative would clearly serve. 3D comparison could probably be effected most easily on the basis of stereo pairs, in either case.

3. Philosophy of LERNIE

The general problem may be summarised as follows: Too many data are accessible only with great difficulty, or not at all, either psysically or because they are unpublished, unclassified or even unordered. In another sense, there is too much published information.

On the otherside, there are too many people spending too much time, energy and money - and suffering too much frustration - trying to (re)assemble the same data, in isolation, in the hope of adding new data.

The overall optimum soluation would appear to be to classify all data straight into banks, on some such lines as suggested by Laflin below, and to make them instantly and globally available in the most acceptable form, as outlined by Wilcock.

At first sight the cost might seem totally prohibitive. The most suitable equipment is priced around £2 million. But such hard costing is easily undermined. First, it needs to be related to the random general background (say, ERNIE's monthly); then, to its specific professional context (say, one year's salary for 1,000) and finally, to the frustration, wastage and loss of potential which cannot be quantified. This becomes particularly forceful when one sets the frustration of those 1,000 - whose salaries, internationally are already being spent anyway - against the comparative wastage of their efforts, on the on hand, and ERNIE's careless rapture on the other.

I see the development of a single, comprehensive, international CPU with core resident systems (a la Scollar) and carefully graded satellite program architecture based on the kind of PLUTARCH blueprints outlined by Wilcock. Its great depth of mental focus will enable it to handle economically a wide range of problems - and graphically - from simple analyses of the kind described by Bradley, and Allsworth Jones' encircled scalograms, to the increasing complexities envisaged by Green and Doran, and in Clark's search for suspected relationships. One of its more fascinating segments will be dedicated to a coninous interactive identification service, fully cumulative, like a simultaneous lightning game of Twenty (thousand!) Questions - developed along the lines of the punched-card key to timber identification. Publication will obviously become obsolete with the advent of instant unit demand feeding and milking of data.

A mirage, you say, and you may be right. But even if you were, it would only be an indicator of the distance and conditions between us and another bit of reality. One tends to forget that somewhere behind every mirage there lies as assemblage of refreshingly hard facts. What we need now is a fellowship or two, to study the lie of the land, and to make a start with that software magic carpet.