

Computers and the biologist

Among biologists, physiologists have inevitably taken most readily to the use of computers; people brought up to know the difference between d.c. and a.c. amplifiers are naturally more inclined to give house-room in their laboratories to further electronic equipment. Yet the past few years have seen the infiltration by computers of most other kinds of research laboratories.

In all kinds of laboratories, the chief use of computers is still in the storage and analysis of data. The principles are as old as computers themselves. Data are digitized, stored within the accessible part of the computer or dumped onto magnetic tape, and analysed by whatever computer program is appropriate. In the 1960s, most large biological laboratories found it expedient to install substantial main-frame computers to provide their members with a data-processing service. Now, however, with the arrival of mini-computers and microprocessors, the trend is towards distributed computer capacity. Individual laboratories are acquiring their own mini-computers, while many pieces of equipment are now designed to operate with their own built-in microprocessors. (Most manufacturers of substantial pieces of equipment, for radioimmunoassay, for example, now provide floppy disks which allow the users to play games such as Star Wars as well as the more sober programs required to analyse the data, no doubt from compassion for those working in laboratories with time on their hands.)

This trend is probably inevitable, and unstoppable. For the user, the obvious advantage is to be fully independent, able to process one's data whenever necessary. Central machinery working on a batch-processing basis is necessarily by comparison an encumbrance. And there has grown up a certain air of disenchantment, at least among biologists and biochemists, with computer systems designed to be accessible by means of computer terminals in the laboratory.

Nevertheless, some of those concerned with the provision of computing facilities in general purpose laboratories are uneasy. Although the new small computers and microprocessors are cheap, the software required for specific specialized function may make serious demands of program-

ming skills. The end result, some gloomy calculations suggest, may be that the central cost to laboratories may be greater than at present. The alternative, that individual researchers should themselves become skilled at programming their own machines, is regarded with apprehension.

Laboratory administrators are also uneasy about the cost of measuring equipment with built-in microprocessors. The convenience and efficiency of these machines is beyond dispute. The suspicion is merely that the value added by the manufacturers of the equipment may not give the user a fair share in the cost benefits of the cheaper hardware now available.

The truth is probably that laboratory administrators are learning a painful lesson in economics. The development cost of software for a commercial measuring instrument is likely to be greater than that of writing a program in a particular laboratory. But the manufacturers of the measuring instruments with built-in microprocessors also find that demand is buoyant, suggesting that people able to afford the new equipment are prepared to pay for its convenience.

There are several well-attested illustrations of the costs that can be saved by home-made equipment. Thus one West Coast neurophysiologist, unable to afford from his current grant the \$20,000 to buy a suitable wave-form analyser, reckons that he was able to produce a better system for a quarter of the cost and at the same time produce a system of analysis better suited to his needs. The underlying dilemma is familiar — to what extent should research laboratories spend their time and energy on the construction of instruments which are nevertheless obtainable commercially.

Not all of the tasks of the general biological laboratory have as yet been delegated to the instrument manufacturers, however. Indeed, even in some of the most familiar tasks, the analysis of electrical output from physiological preparations for example, formidable intellectual difficulties keep cropping up.

The standard record in such experiments consists of a voltage (or several voltages) varying with time, on a time scale of milliseconds or thereabouts. Given a visual trace of such a voltage, it is possible with a ruler to estimate parameters such as the peak voltage, the time taken to reach that voltage and so on. Computer systems store the same information in digital form, and there are constant (if groundless) worries about the extent to which such data embodies "everything". More seriously, physiologists are likely to insist that some form of visual record should be obtained in parallel with whatever computerized analysis is found convenient. Machines may be able to analyse complicated sets of data once they have been provided with criteria for setting about this task, but

cannot recognize novel kinds of patterns.

So much is evident in the difficulties which keep cropping up in the research programmes intended to provide more sophisticated monitoring of the physiological condition of hospital patients, people in intensive care units for example. Many of the systems developed for this purpose depend on the recognition of abnormal trends in, for example, heart-beat rate, or variation from one cycle to another. One difficulty is that too little is known about the statistical properties of normal heart-beats to provide a sure yardstick for abnormality. The result is that false warnings abound, while few are confident that the systems now in use are capable of recognizing more than a small proportion of impending crises.

But what of the application of computers to tasks in biological research that have hitherto been beyond the scope of conventional techniques?

Many of the innovations in the past few years are comparable with those in other fields. It is unthinkable that X-ray diffraction data should, these days, be analysed by the traditional methods, and computer techniques are used as a matter of routine in construction of the models of most molecules, biological or otherwise. Quite apart from the way in which computing machinery has given people the courage to tackle problems that would otherwise have been neglected, the more economical use of data has made it possible to work with lower X-ray intensities.

Another development of this kind is in the construction of molecular models. Modelling biological molecules of any complexity is a formidable task, and this is the principal incentive for the development in the past decade or so of computer programs which can handle such tasks quickly but, more important, can in principle explore all possible configurations of a complicated molecule.

These are essentially data-handling tasks. The most recent development of this kind has been prompted by the sheer size of the nucleotide sequences now being accumulated. It is natural that these should be stored in computer form, and plans for the development of such a computer bank have been discussed between the National Institutes of Health in the United States and the European Molecular Biology Organization. One obvious elaboration of such a system is the use of computer programs for searching nucleotide sequences for particular configurations of nucleotides, for example for configurations that define the points of action of particular restriction enzymes. But the construction of a complete nucleotide sequence is not always straightforward, given that it may have to be inferred from the nucleotide sequences of overlapping fragments of a complete DNA molecule, which is why there are now ambitions to solve such essentially combinatorial problems by using computing machinery. □

Computer model of the protein crambin

