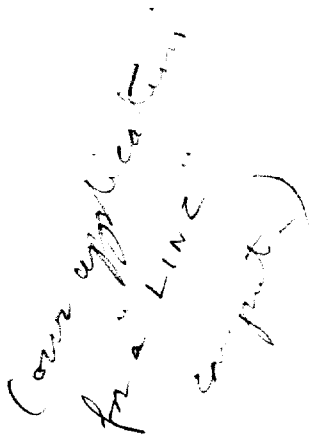
An Instrumentation Crisis in Biology

Joshua Lederberg
Professor of Genetics
Stanford University Medical School
Palo Alto, California

May, 1963



AN INSTRUMENTATION CRISIS IN BIOLOGY

Physiologists, especially students of the nervous system, have a long history of involvement with electronic instrumentation, many of the primary signals being already in electrical form. Biochemistry has made much less effective use of such instrumentation - laboratories like Britton Chance's being quite exceptional. In consequence, we face some of the most profound issues of biology, in molecular biology the sensitive detection of macro-molecules, and the specification of their ultrastructure and metabolism, with tools that are astonishingly primitive by the standards of instrumentation available in other fields. This is especially true of the range of optical instruments, which are slow, imprecise and require inordinately large samples of material compared to some plausibly attainable possibilities. A second area of data processing in which a great burden of manual effort can be lifted is particle-counting which, in one form or another, is the fundamental measurement in many aspects of biology (especially microbiology, cytology).

Finally, perhaps most exciting, an immense amount of information is still locked up in spectra (optical absorption, magnetic resonance, rotary dispersion, mass spectra) and similar fingerprints, which require the intensive development of the "man-computer symbiosis" for adequate resolution. While such successes as the structural analysis of myoglobin have had well-deserved attention, even they focus attention on the importance of further instrumentation development to solve such problems on a broader scale (each one should not be a tour-de-force, at least not after the first one.)

The inadequacy of current art in biochemical instrumentation was brought home to us in our efforts to meet the mission requirements of exobiological studies, but they are equally pertinent to present efforts in the terrestrial biochemical and microbiological laboratory, except of course we at least do have the traditional tools to work with meanwhile. We have been delighted at the opportunity created by NASA's programmatic and financial interest to try to contribute conjointly to exobiological and terrestrial needs in biological instrumentation.

Over the past three years, we have gradually been organizing an instrumentation group within the Genetics Department. Under Dr. Levinthal's immediate direction, its scope can be indicated by a present staff of three additional electrical engineers, (L. Hundley, H. Horn and N. Veizades), a physicist (M. Mandel), and for the exobiological work, a biochemist, (Dr. E Shneour). Two or three additional appointments are in prospect.

Some of the more directly mission-oriented work has been reasonably successful, for example the mechanical design and prototype construction of a "multivator", a compact laboratory for sampling and chemical analysis of Martian surface dust. On the other hand, as any experienced hand would have predicted, many of the "bright ideas" we have developed, either as requirements for instruments, or as their designs, have bogged down

when the construction and debugging of the devices took months instead of weeks. The worst result is perhaps the loss of interest in the original purpose of the design; the next worst is the reaction of hypercaution in deciding whether to go ahead with a given project.

The Instrumentation Crisis is thus deepened by the lack of flexibility that these considerations imply, and the practical attitude of disdain for preoccupation with instrumentation displayed by most members of the biochemical scientific fraternity.

Digital computation may help answer these needs in several ways. Precision can be improved in data-processing links as an inherent virtue of a digital system. Unfortunately the original data are generally in analog form, and the accuracy of analog-to-digital conversion will be a limiting factor. However, precision is also very often a signal-to-noise problem, and an ideal instrument should have the flexibility to allow accuracy to be purchased at the price of speed, in accordance with local needs - the memory capacity of the computer for averaging over time, and the use of correlation techniques to extract signals from noisy inputs, suggests the application of simple computer techniques to improve the utility (and to simplify the potential design) of such workaday instruments as the absorption spectrophotometer. Since many problems cannot escape the dilemma of requiring measurements of differences between larger values, this is no luxury. Probably more important is the construction of (at least) the prototypes of new instruments by programming a general-purpose computer to set up the control and signal-processing systems, instead of de novo construction.

Further, mechanical corrections to maintain linearity and stability play a large part in the design of most instruments - for example the slit-width control to maintain constant reference brightness with varying wavelength in a spectrophotometer, or the slide-wire bridge to obtain an intensity ratio in its output recorder, and these generally impose severe penalties in the complexity of design as well as the accuracy of the instrument. A fairly small memory and computing element would replace most of these expedients. It is likely, but not certain, that the specialized training needed to make some intelligent use of <u>programming</u> in instrument design is less than needed for hardware construction, a feature that would make innovation more widely accessible to other biochemists in a university environment. The computer is, of course, a much more complicated instrument than any of these, but it is intrinsically more accurate; it has programmable versatility, and most important it was designed and built once and for all.

Finally the course of <u>future</u> developments certainly seems to call for the preeminence of digital (or mixed) computation, and the emergence of more and more compact and reasonably priced systems.

We have therefore determined to reorient our entire program towards the most efficient use of digital computers as on-line elements. Since our resources would still not justify capturing a larger machine, our opportunity to proceed along these lines have only recently opened up through two possibilities: (1) a time-sharing system, and (2) the LINC.

We do not know which of those affords the greater promise--probably each of them has a particular place, and we can hardly ignore the possibility of further interaction between them. Our special attributes for the evaluation of LINC then perhaps include (a) newcomers to extensive use of digital computation; we have no on-line experience at all, except insofar as some use of digital elements has appeared in our own instruments; (b) a team effort, intended to serve a variety of instrument needs, but with special emphasis on biochemical analysis (nucleic acids, proteins, enzyme assays); (c) a comparison of utility of LINC and of time-sharing in a variety of situations.

The time-sharing system referred to is the special interest of Professor John McCarthy of the Stanford Faculty in Computer Sciences. Our laboratory will participate in a program he is designing to allow some few stations to share access to a PDP-1 computer, which will in turn be coupled to the IBM 7090 system now in operation. The PDP-1 is scheduled for delivery next month; within a few months thereafter we should have telephone or direct wire access to it from our own laboratory. The final configuration remains to be worked out; some buffering to mitigate the disturbance of interrupts is intended to make the system useful on line.

The current active projects reflect an over-conservative reaction to recent problems; I am, therefore, also appending a list of other items we would keenly like to be able to manage, and urgently hope to get into. Some tentative priority is given, but this will be tempered by the acquisition of some additional people, and by a reappraisal of the actual utility of LINC or the PDP-1 time-sharing arrangements in each case.

CURRENT PROJECTS: ACTIVE CONSTRUCTION AND TESTING

Some additional details are given in appended descriptions:

Multivator. A device for the acquisition of samples of planetary dust, their distribution to individual compartments, programmed release of solvents and reagents, and the determination of enzymatic activity or growth by various photometric measurements (absorption, fluorescence, scattering, polarization, scintillation).

Mark III incorporates a flying-spot scanner to identify local areas of enzymogenic fluorescence, the local illumination enhancing signal/noise. The signal is expected to reflect localized incidence of microorganisms; the noise comes largely from the background fluorescence of the soil sample eluate and from spontaneous degradation of the fluorogenic substrates (e.g. fluorescein phosphate).

Videoscan Spectrometer. A fast scanning spectrophotometer based on the projection of a spectrogram on the target of a signal-storing image-forming television tube. This was primarily intended for microspectrometry as a means of searching for microbes in soil; this application will require the delivery of newly developed, UV-sensitive video tubes to complement Zeiss ultrafluor optics obtained with Caspersson's cooperation. The basic system has been completed, and is being tested by being installed in a model E analytical ultracentrifuge to allow online reading of absorption profiles (hopefully also absorption spectra at each stratum) of DNA in the course of pycnographic fractionation. As these centrifuge runs may require two to several days, the accurate determination of the profiles online should greatly shorten machine time; as the signals are already in electrical form, the computation of peaks, band decomposition, integration and linearity corrections, would all be facilitated.

The same video system can be and has been used for more general purposes - mensitometry on photographic spectrograms, wherein it is being compared with flying-spot video. It is also being set up for simplified frame-differencing, to allow the discrimination of moving targets: in our application, the tracks of microbes with purposeful motion. Having spent some years in "human" microscopy of individual bacteria, in connection with the biochemical genetics of their flagella, I am looking forward to the possibility of computer-assisted analysis of this primitive behavior.

The control system of the videoscan is a versatile one, especially to allow alternate scanning of a reference and sample field; then a line-interval delay allows the superimposition and comparison of sample/reference, giving an equivalent dual-beam mode. Plainly, most of this could have been quickly by-passed by exploiting a general purpose computer.

Colony Counter (Iconumerator). The most tedious operation in bacterial genetics is counting colonies of plates. We inherited an Iconumerator (a flying-spot pulse analyzer) of World War II vintage from the DuMont Laboratories

Colony Counter (cont)

and have been trying to make it work properly. This is difficult, apart from the limitations in its one-line logic.

The most serious problem is the reliability and resolution of the delay line, which stores information for common-mode rejection of a colony count at second encounter with the scan beam. We have been speculating about a major redesign and rebuilding with modern components; would very much prefer the course of simulating a new design and trying it out quite extensively before making a large commitment of time, effort and money. Quite possibly the whole issue can be evaded by other mechanical arrangements (a microbiological machine that disposes the inoculum in a linear array, so that counting can be done by counting pulses along a single axis).

Curve Reader and Analyzer. This is an elementary combination of XY-plotter and curve follower, and tape-recorder, synchronized to a small analog computer. Chart records are thus translated to signals on each of several channels, which can then be fed synchronously through the analog computer for analysis. It has been useful for elementary functional transformations, e.g. of recorded and published absorption spectra, and for instructional purposes.

POTENTIAL PROJECTS: ACTIVE CONSIDERATION, especially for digital enhancement

1. Precision Photometry; laboratory measurements

The straightforward enhancement of the performance of existing instruments and techniques may be the most elementary but rewarding application of computer techniques, using the principles already summarized. Spectrophotometry plays a key role in all our work on bacterial DNA; we would like to be able to measure more accurately in the range below o.1 optical density unit. Even a very simple instrument in which reference information can be sampled and the signal time-averaged over some interval should be able to do this for us.

An analogous statement can be made for spectrofluorometry. Further, Dr. Lubert Stryer is soon joining the Biochemistry Department, and is especially interested in cooperating here towards the development of new analytical techniques which require the utmost detectivity in fluorescence measurements.

2. Depolarization of fluorescence; measurement of relaxation times. One of the most promising simple methods for estimating molecular size of DNA and following it through thermal transitions is the measurement of relaxation times after electric orientation of the solution. For sensitive detection, we propose the use of fluors coupled to the polymer, and the use of polarized exciting light.

As relaxation times of hundreds of microseconds are anticipated, the entire control and analysis program should be compatible with the LINC's capabilities. This approach may have special promise for the detection of macromolecules in complex solutions for exobiological purposes, among others.

3. Optical rotatory dispersion plays a special role in long-range thinking about exobial detection for reasons we owe to Pasteur. Precision measurements, especially of highly absorbing materials like polynucleotides, are essentially a problem of extracting signal from noise. Our colleague, Professor Djerassi in the Chemistry Department, is especially interested in such instrumentation, and is a well-known authority on its utilization in chemical analysis.

4. Microscan mass spectrometry. This is the most ambitious project but may be the most hopeful in meeting urgent requirements at the focus of several interests in the department; genetic chemistry, neurobiology, and exobiology.

We propose to scan a specimen at (crude) electron microscope resolution with a high intensity beam along the lines of the electron probe microanalyzer. However, the scanned spot would be volatalized and the gas fed to the beam of a fast mass spectrometer (e.g. - the Bendix Time-of flight instrument) for mass analysis. At low mass numbers, and with

Microscan mass spectrometry (cont.)

various other tricks, it might be most useful to localize tritium at resolutions, perhaps even sensitivities, exceeding current autoradiography. (The LINC would play a special role in processing the spectral outputs for tape storage, selection and readback, either through the LINC or through another computer for even more detailed analysis. It could, of course, also furnish the principal control mechanisms.)

5. Microfluorometry. Dr. Boris Rotman, in this laboratory, devised a method for the detection of enzyme activity not only for single bacteria, but even single enzyme molecules, based on the microfluorometry of fluorescein released from a galactoside conjugate. He has since moved to another laboratory in the area, but we are maintaining a close collaboration. Much of this work requires a very laborious sizing of microdrops under the microscope, and then a determination of fluorescence, both of which should not be too difficult to instrument for mechanical operation.

As the analytical technique is mainly based on Poisson statistics, a large number of drops must be analysed, not to infinitesimal precision. This equipment would be of very great value to both of us in studying the role of DNA in instructing the cellular synthesis of enzymes.

ADDENDUM

Further extensions of computer technique in human biology and genetics.

In addition to laboratory experiments these studies often require the analysis of large volumes of vital statistics and similar data. The indispensability of computers in tabulating information from such files is of course well understood. However, there are serious problems in the accessibility of computed data in this area, just as there are in experimental work. As important as are the obvious current applications of computers in data processing, we should not be content with the present system but work for more flexible means of experimental access to such data through "real time computation".

The small computer, like the LINC, or the time-sharing system, furnishes an answer to this need. Very extensive data files (up to 300,000 tab cards) can be stored on a single reel of magnetic tape which need require no more than five minutes to pass through the computer. A flexible system would allow this to be done under direct manual control and with immediate console display of computation on these data files. In most applications, except those that involve very extensive sorting and resorting, the actual computing time in a large computer like the IBM 7090 is relatively short compared to the time required for the input-output processes. With the appropriate organization of the computer facility, it should therefore be entirely feasible for the investigator to deal directly with the large data file in real time, while his attention is still concentrated on the problem at hand and in the formulation of new hypotheses for prompt testing. Only in this way can we really make the most effective use of the combination of human ingenuity and constructive imagination and the computer's capacity to undertake a humanly impossible task of drudgery in computation.

The most evident application of this approach in genetics should be in the analysis of vital statistics. In the experimental sciences the nearly analogous applications will involve searching through data accumulations like spectra and other physical properties, the experimenter forming generalizing hypothesis and using the computer to test them against the file.

Experienced insight into the relative roles of the human and machine components of these systems should be invaluable in the further development of artificial intelligence, the programming of machines to simulate as far as possible those human cognative processes that we can begin to understand. These generalizations of human intellectual capability, and their delegation to machines, continue to refine the strategic role of the human element and to give it increasing leverage in the solution of complex problems.