methods for the first time in the extraction of dynamic information from pyelograms. In this way, governing parameters determining the stability of a unidirectional flow towards the bladder and the prevention of ureteral reflux can be studied and properly evaluated. Furthermore, reliable quantitative documentation for developing disease states will be possible by the integration of the measurements of the outlines and significant features of radiographs with other archival medical records. In this way, it will be possible to document reliably the time dependent changes in the upper urinary tract dynamics of patients with spinal cord injuries, recurrent urinary tract infections, congenital abnormalities, stones, etc. The direct result of this classification is to reduce to a manageable level the amount of repetitive work and provide the urologist with a direct method of realizing the diagnostic information asked without having to handle massive loads of films and data.

(d)   Methods and Procedures.

A great effort has been devoted during the last few years to image evaluation and enhancement. The group at the Jet Propulsion Laboratory has continued its work on image enhancement with application to space television and biomedical imagery. For the purposes of this study, it is not our desire to repeat any of this pioneering work. The field would obviously provide us with newer, better, and more efficient algorithms for picture encoding, computer processing, and pictorial pattern recognition. Philosophically, our approach will be in the domain of feature extraction algorithms which will outline the renal, ureteral, and bladder projections from each frame in the most economical way, avoiding brute force methods.

Additional features of interest will include video densitometry within outlined structures related to the concentration and thickness of radiopaque material. This information can be combined with peripheral boundary measurements to infer relative cross-section information.

The time evolution of pyelograms includes considerable redundancy and correlation from frame to frame. This information can be used to model local system dynamics (frame to frame variations) and hence direct the data processing. In this way the relatively sophisticated pattern analysis algorithms for edge detection, structure skeletonizing, and densitometry can concentrate on small important subsets of the large amount of raw data involved. We can minimize computing and time resources required by utilizing these solution guided processing schemes.

The initial orientation of the computer to an image sequence may require human interaction as well as to assist in processing difficult frames. We expect the application of these results to succeeding frames as described above to minimize the need for human interactions.

Algorithms which may be used for automated computer analysis of the images include dynamic threshold and maximum gradient techniques for edge detection. Continuity conditions can be imposed for sequential

edge following. Techniques similar to these are being successfully developed for karyo typing and arterioangiogram analysis. Relative video densitometric information can be used to determine ureter size by measuring variations in density about a local mean and relating these to cross sectional area. This is accomplished by correlating mean density to mean boundary separation and assuming vessel symmetries.

Time evolutions of extracted information will be studied through various on-line graphical presentations. Deviations from predicted local model behavior may be used to identify time dependent anomalies which could be of interest.

The pyelogram information is collected on 35 mm film and significant archives of patient histories exist and will be studied. Initially, to develop techniques, we will use existing video scanning equipment at Stanford (SLAC) to digitize the entire frame sequences of interest. In order to conveniently store this large volume of data for computer processing we may use data compression techniques which take into account intraframe as well as interframe correlations. A scheme requiring relatively little computational effort to encode and reconstruct the video information would be a Huffman code built around picture element difference statistics.

In the longer term we want to obtain resources to build a computer directed scanner so that only the pertinent information need be measured from the film.

(e) <u>Collaborative Affiliations.</u>

A continuing dialogue and exchange of ideas in both the experimental and theoretical aspects of image processing will be maintained with Dr. Harrison's group. Although the contrast and time constants of cardiac catheterization is very different from pyelography, it is hoped that this cooperation would prove economically and scientifically beneficial.

(f) <u>Significance.</u>

The development of computer based characterizations of radiographs would be of practical significance to the present and future need of urologists and radiologists. The information accrued from such a system would span a broad range of benefits in specific areas of quantification, standardization, and information storage of otherwise neglected parameters in pyelography.

The potential future development of pattern recognition methodologies which can be realized from these evaluations, can handle more variability and allow the consideration of biomedical applications. It is realized however, by other workers and us that instead of trying for completely automated systems, we would use an interactive man-machine system that would allow human intervention in questions and tasks not easily automated. Thus, a number of interesting and difficult applications for computer assisted analysis of radiographs can become technically feasible.

(3) Automated Gas Chromatography/Mass Spectrometry Analysis
     Prepared by Tom Rindfleisch

(a) Introduction

This section of the proposal is concerned with the design and development
of the computer hardware and software components necessary for a fully auto-
mated gas chromatography/mass spectrometry (GC/MS) system. This work repre-
sents an extension of a portion of the existing DENDRAL research grant
(NIH grant RR-00612). Significant progress is being made in all phases of
the DENDRAL research as summarized in the succeeding sections. We have
gained a greater experience with the data system requirements for the auto-
mated collection and analysis of mass spectrometry data as well as the
limitations of the existing ACME system for developing such system. Because
of the critical requirements for information integrity throughout an automated
GC/MS system for medical applications and because the DENDRAL project is a
natural environment in which to explore "intelligent" information handling and
instrument control problems, we are proposing this more extensive automation
effort. This work complements the on-going DENDRAL artificial intelligence,
chemistry, and instrumentation research as explained in subsequent paragraphs.

(b) Problem Statement

The combination of gas chromatography with mass spectrometry (GC/MS) has had
a tremendous impact on analytical problems in organic chemistry and bio-
chemistry over the past decade (1). Increasingly GC/MS is being used as a
tool in clinical and medical research applications to identify metabolites
and other materials contained in body fluids. For example, Biemann and
collaborators (2) describe a dramatic series of events following the admis-
sion of an unconscious patient to a hospital following a drug overdose.
GC/MS analysis provided an identification of the drug used in this suicide
attempt. Fales and Milne have been active in the identification of abused
drugs. They describe (3) the use of GC/MS for the analysis of drugs separa-
ted from the stomach contents of 45 would-be suicides, and the resulting aid
to proper treatment of the patients involved.

Medical research applications of computerized GC/MS have included the detec-
tion of metabolic disorders of genetic origin from an analysis of the organic
chemical constituents of a patient's body fluids (usually blood or urine).
Jellum and his associates (4) in Norway have been active in this field and
through largely manual methods have been able to identify four previously
undescribed metabolic diseases of genetic origin based on urine analysis.

These examples serve to illustrate the potential of GC/MS analysis as a
tool in medical research and clinical applications. The basic power of the
techniques lies in the ability of the gas chromatograph to physically separ-
ate and pass the components of a complex mixture into a mass spectrometer.
The spectrometer makes measurements leading to a "fingerprint" for identi-
fying each component of the mixture. In urine samples such as studied by
Jellum, et al., such a mixture may contain several hundred components and
may be subjected to fractionation prior to GC/MS analysis. The amount of

data contained in the output of such a GC/MS experiment is very large and the procedures for extracting significant information including interpretation are complex. An experiment may last for 2 hours with a spectrum containing 10^5 samples of data collected every 5 to 10 seconds. Out of the possible ensemble of 10^8 data points must come an identification of all components with as little ambiguity as possible. Computer based data systems are essential for any effective utilization of this powerful technique.

Several computer-based approaches have been utilized to aid in the analysis of low resolution mass spectral data, whether or not collected by GC/MS (5) systems. For materials to be identified within a known class of possibilities, the potentials of library search routines have been explored (6). These procedures are frequently ambiguous as they use only a subset (low resolution spectra) of the information which a mass spectrometer can provide. Furthermore, in many medical research situations it is precisely the unexpected or previously unknown materials which are of greatest interest (4). Such problems cannot be solved within the domain of a library and currently fall back on human intervention to resolve ambiguities or synthesize new solutions. This limitation restricts considerably the utility of GC/MS systems because of the effort and time delays required to explore new situations.

The on-going DENDRAL work (7a-i) at Stanford (RR-00612) offers a far-reaching solution to this problem by designing into computer programs the ability to construct explanations for mass spectra in terms of chemical structure. Recently, these efforts in the context of high resolution mass spectra, have had considerable success in dealing with estrogenic steroids (7h). Future work will expand these capabilities to more classes of compounds as well as generalize the heuristic rule-forming processes to allow automatic computer extensions.

Manual, as well as automated interpretive procedures (7c,d,h) can utilize a variety of ancillary information (e.g., high resolution, low ionizing voltage, metastable data, and NMR data) to produce reliable and unambiguous results. Present low resolution GC/MS systems, because of limitations in system and/or instrument designs, are generally incapable of collecting all possible information on which to base an analysis during the finite interval of a gas chromatograph effluent peak. A realtime selection of instrument mode and information optimization type and quality is required.

Acquisition and reduction of mass spectral data in realtime have progressed to the stage where automation and closed-loop control are feasible and desirable. Closed loop automation of data extraction and instrument control processes places additional burdens for capability, integrity, and reliability on the overall data system. Because of the complexity of data reduction and interpretive processing, great care must be taken throughout the system to avoid the destruction of or artifactual invention of information.

The objectives of this portion of the proposal are to develop and demonstrate a fully automated gas chromatograph/mass spectrometer system in collaboration with on-going DENDRAL chemistry, artificial intelligence, and instrumentation research.

Specific objectives of the research include:

1. The development of autonomous and reliable instrument control and information extraction programs capable of reacting to a hierarchy of GC/MS response requirements within the context of a time-shared computer system.

2. The integration of evolving DENDRAL artificial intelligence programs to interpret extracted information and to provide feedback within the system specifying information requirements to insure optimal sample management.

3. The application of the developed system in cooperation with collaborating chemists and medical researchers to problems in the analysis of steroids and other metabolites found in biological fluids.

(c) Background

The elucidation of the systematics by which chemical compounds fragment under electron bombardment has a large literature with very significant contributions from the laboratory of Professor Carl Djerassi at Stanford. These systematics form the basis for computer automation of the interpretation of mass spectra – the DENDRAL project.

DENDRAL is a set of computer programs which have developed over a period of several years, initially for the interpretation of the low resolution mass spectra of specified classes (ketones, ethers, amines, alcohols, thiols, and thioethers) of alicyclic compounds (7b-7h). Subsequently, this theory was extended to include the high resolution mass spectra of the estrogen class (female sex hormones) of steroids (7i). This program has also demonstrated its ability to identify the components present in laboratory-made mixtures of estrogens. At the present time work is progressing with crude estrogenic mixtures obtained from biological sources. The successful completion of this project will represent a new, rapid approach for the identification of estrogenic steroids without the necessity of first derivatizing and then analysing the mixture by gas chromatography. Heuristic DENDRAL is also being enlarged to accommodate a theory of mass spectrometric fragmentation of other classes of steroids and alkaloids, utilizing high resolution mass spectra.

Metastable ions formed in the first field-free region of a double focussing mass spectrometer (so called defocussed metastable ions) have been used by mass spectroscopists for the identification of parent-daughter ion relationships. As part of its spectrometry theory, Heuristic DENDRAL will use this additional type of experimental data. A recent paper (8) described a new use of defocussed metastable ions for the unraveling of competing fragmentation pathways.

Meta-DENDRAL research efforts are aimed at the computer formulation of scientific theories based on the examination of related sets of data. Such

a capability will allow the automatic extension of computer capabilities for mass spectrum interpretation by the inference of new rules. To date these programs are capable of writing primitive rules about fragmentations and the influences of molecular parameters (e.g. substituent effects).

> Work of others in the field. Data systems which can cope with the large accumulation of spectral information generated during a low resolution GC/MS run have been developed in a number of locations including Stanford (9). These systems run open loop in that they systematically collect data from sequential low resolution spectrometer scans, reduce the data based on instrument calibrations, and provide the chemist with the ability to retrieve particular spectra corresponding to gas chromatograph effluent activity. However, even when coupled with library search procedures (6), there systems make few, if any, intelligent decisions about the data, and provide the chemist with few clues about the validity of his results. Our approach will be to develop techniques to validate results under closed loop control based on instrument performance parameters and ancillary information such as the routine use of high resolution data. We will make use of the work by others in the development of suitable library search routines.

(d)   Rationale

The power of the gas chromatograph/mass spectrometer data system as a medical research tool derives from the ability of the gas chromatograph to physically separate microgram quantities of complex mixtures followed by the mass spectrometer to identify each constituent from its "fingerprint" mass spectrum.

For each of the gas chromatographic peaks, the basic function of the mass spectrometer is to ionize sample molecules which then fragment and, through electromagnetic separation, to measure the abundance of fragments with different masses. At high resolution the elemental composition of the various fragments can be determined. These abundances are related to the molecular structure of the sample material and these relationships can be used by inference to derive the structures for unknown sample materials from their mass spectra. There are numerous modes of operation of a mass spectrometer which allow the measurement of ion abundances with varying time, mass, resolution, and ionization energy, as well as enable the observation of delayed or metastable ion fragmentation pathways. Not all information in all modes of operation can be collected during a gas chromatographic peak because of limitations in data rates, instrument sensitivity, and sample flow into the ion source. Conversely not all collectable information is necessary for the identification of an unknown. The optimum experimental conditions producing the most relevant information in the shortest time are not predictable a priori for an unknown material. Thus closed loop computer analysis of the spectrometer with subsequent feedback control of its operation could maximize collected data quality and ensure the collection of needed information for the interpretation of an unknown structure.

The essence of our proposal is to design the necessary information handling and system control intelligence to complement the DENDRAL spectrum interpre-

tation intelligence and to integrate these elements into a reliable, autonomous GC/MS system. The core of this work will be to design programs at the various processing stages shown in Figure 1 which allow the system to perform the required functions. The time constants and data rates involved in the various aspects of a typical GC/MS experiment range from ~1 msec to ~10 seconds. These requirements are based on the typical 5 - 30 second duration of sample uniformity in gas chromatographic peaks. The logical sequencing of the loop element operations of Figure 1 is dependent upon the sequence with which information becomes available and the degree of overlap possible between successive operations. Figure 2 shows conceptually how this sequence takes place.

The GC/MS data systems existing today run almost entirely open loop in that there is no attempt to modify experiment execution based on extracted results. The processes involved are implemented with inadaptive algorithms so that if instrument performance, or data quality do not fall within parameter specifications, information may be destroyed, ineffective filtering may occur, or catastrophic system failures may result. Specific examples of where added intelligence is required exist throughout the system:

(1) Failsafe data collection and management capabilities must be built into the system to accommodate the inherently variable data length and peak arrival rate statistics of the spectrometer output.

(2) Reliable and fast methods must be developed for detecting and resolving overlapping peaks in the gas chromatograph and mass spectrometer sensor outputs.

(3) The quality of extracted information must be evaluated based on instrument performance parameters and ion statistics.

(4) The characterization of instrument performance parameters must be continually updated and evaluated to allow optimum control of parameter settings for scan, focus, resolution, source and reference pressures, etc.

(5) The successive operations on extracted information must adapt their behavior based on the character and quality of their input information and must add their effect on uncertainties at their output.

(6) The overall management of resource allocation must be based on priorities derived from the on-going problem solving and interpretative processing to maximize the effectiveness of applied resources, to decrease processing time, and to minimize computing cycle consumption.

The evolution and collaborative application of this system will be complementary in nature. The conception, design, and implementation of the system benefit from experimenting with its applications. Conversely the power of the automated system allows the systematic exploration of new areas in medical and chemical research.
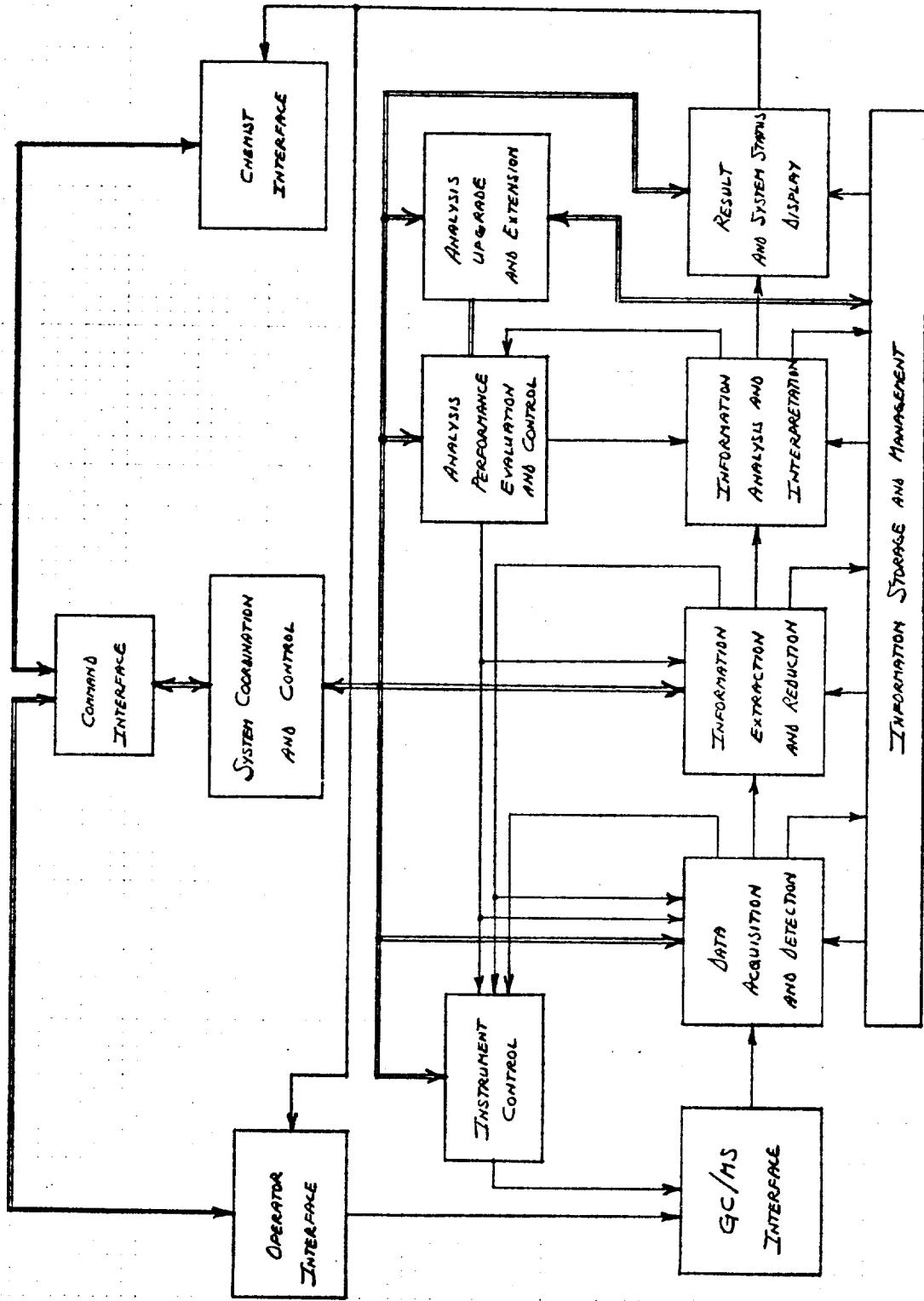
Figure DENDRAL-1

Real Time Mass Spectrum Interpretation System

TLR 2/17/72
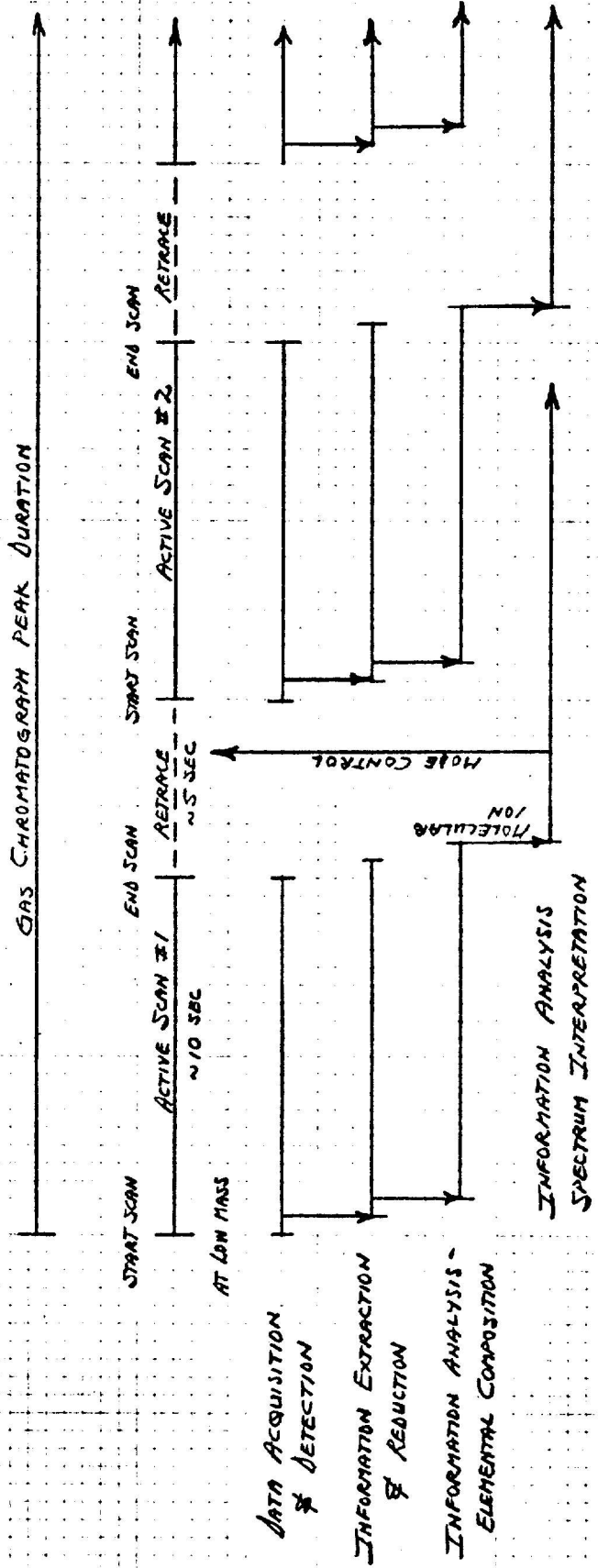
Figure DENDRAL-2

(e) Methods and Procedure

The implementation of the proposed automated gas chromatograph/mass spectrometer system will be a highly collaborative effort drawing as much as possible upon capabilities existing in laboratories at Stanford and elsewhere. Specifically we will use:

(1) The GC/MS-computer instrumentation and data system interfaces existing and being developed at Stanford under DENDRAL and NASA grants.

(2) Appropriate modifications of existing library search algorithms and data bases to effectively utilize (3) below.

(3) The artificial intelligence programs for mass spectrum interpretation being developed under the DENDRAL and ARPA grants.

In addition to these collaborative interactions, we will draw heavily on the satellite machine support capabilities, extended realtime system functions, and PDP-10/satellite hardware and software systems being developed under other sections of this grant application.

The various elements required in an automated GC/MS system are shown in Figure 1. These elements perform a variety of functions including:

(1) Data acquisition and detection. This element accepts the raw spectrometer output and detects peak information above a dynamic background threshold. Based on peak arrival statistics, control of the spectrometer scan may be used to increase ion collection efficiency (this would be based on similar work being done by McLafferty, private communication).

(2) Information extraction and reduction. This element extracts separated peak amplitude and position information from raw data. Instrument calibration data are used and resulting data quality is estimated. These quality measures can optimize instrument parameters.

(3) Information analysis and interpretation. This element computes elemental compositions as required and applies library search routine with appropriate verifications based on spectrum predictor routines. If no solution is found more basic DENDRAL theory construction routines are used to identify the unknown.

(4) Analysis performance evaluation and control. This element directs the search for an explanation of the sample spectrum using available a priori information. When ambiguities arise, control information is directed to obtain appropriate additional data.

(5) <u>Analysis upgrade and extension</u>. When new solutions outside of existing system capabilities are encountered this loop element incorporates these extensions into the system. Such extensions may come from META-DENDRAL analysis or from chemists.

(6) <u>Result and system status display</u>. This loop element provides the system user with rapid volatile plots and displays of on-going experiment results and status. Hard copy is available off-line.

(7) <u>Instrument control</u>. This loop element coordinates and implements control requests on instrument performance such as parameter adjustment or mode change by planning and issuing the appropriate electronic commands.

(8) <u>System coordination and control</u>. This loop element receives and maintains status and performance data relating to various system elements and guarantees the appropriate sequencing of interdependent operations. This element also coordinates system operation changes commanded from the outside.

(9) <u>Command interface</u>. This loop element decodes commands and control information received through the instrument operator or chemist user interface.

(10) <u>Information storage and management</u>. This element includes the organization and storage of spectral information and the ability to access this data on demand.

Models by which the computer can assess and optimize its performance will be developed based on physical principles for instrument performance and heuristic schemes for control and interpretation protocols.

Instrument control functions will be implemented as appropriate for parameters such as gas chromatograph temperature programming and mass spectrometer scan control, mode selection, resolution control, scan dwell, etc. The coordination of these parameters in terms of automated setting determination and sequences for control implementation will be developed heuristically from models of instrument performance and operator procedures.

Our concept of overall software organization follows the functional information flow shown in Figure 1 coupled with the timing interactions shown in Figure 2. The shortest term response requirements (~1 msec.) exist for the data acquisition functions and will be implemented in a dedicated machine interfaced to the GC/MS. This machine also allows open loop operation of the instruments in the existing modes during development of the integrated closed loop system. The other elements of the system will be implemented as subprocessors in the PDP-10/satellite extended realtime system affording required response without total commitment of the PDP-10 system.

(f) <u>Significance</u>. Low resolution GC/MS has become one of the most widely used and most powerful techniques available to the organic or biochemist (1). The potential applications of these techniques in medical research and practically in the clinic have just begun to be explored (4). Closed-loop control of this instrumentation would permit rapid exhaustive analysis of tissue extracts across large populations of individuals in various medical contexts and may provide new discoveries important to public health.

Extension of GC/MS to routine operation of the mass spectrometer at high resolving power would be an important breakthrough in terms of the specificity of information available per microgram of sample, compared to low resolution techniques.

The integration of library search techniques with the screening power of a spectrum predictor and the analytical capabilities of Heuristic DENDRAL would provide a powerful data analysis capability which would exploit the advantages of each approach.

These techniques are of unique importance to medical science since they alone of the current physical methods have sufficient sensitivity and analytical precision to study human biochemistry at the molecular level.

(g) <u>Facilities Available</u>. The research in this proposal will draw heavily upon the PDP-10/satellite computing resource we are proposing to establish. We have available two gas chromatograph/mass spectrometer systems which will be involved in this research including a Finnigan quadrupole instrument in the Department of Genetics and a Varian-MAT 711 instrument in the Department of Chemistry. Also available in the Department of Chemistry are MS-9 and Varian-MAT CH-4 instruments.

(h) <u>Collaborative Arrangements</u>. The proposed research project is a highly interdisciplinary effort involving collaboration between Professor J. Lederberg (Department of Genetics), Professor C. Djerassi (Department of Chemistry), Professor E. Feigenbaum (Department of Computer Science), Dr. B. Buchanan (Computer Science), Dr. A. Duffield (Genetics and Chemistry), Dr. D. Smith (Chemistry), and the Instrumentation Research Laboratory. The proximity of these people and facilities offers a highly unique opportunity for collaborative interaction.

REFERENCES

1. For pertinent reviews see: C. G. Hammar, B. Holmstedt, J. E. Lindgren
   and R. Tham, Advan. Pharma Col. Chemother., 7, 53, (1969); J. A. Vollmin
   and M. Muller, Enzymol. Biol. Clin., 10, 458 (1969)

2. J. R. Althans, K. Biemann, J. Biller, P. F. Donaghue, D. A. Evans,
   H. J. Forster, H. S. Hertz, C. E. Hignite, R. C. Murphy, G. Petrie and
   V. Reinhold, Experientia, 26, 714 (1970).

3. H. Fales, G. Milne and N. Law, reported in Medical World News, February
   19, 1971.

4. E. Jellum, O. Stokke and L. Eldjarn, The Scandinavian Journal of Clinical
   and Laboratory Investigation, 27, 273 (1971).

5. A. L. Burlingame and G. A. Johanson, Anal. Chem., 44, 337R (1972).

6. H. S. Hertz, R. A. Hites and K. Biemann, Analytical Chemistry, 43, 681
   (1971), S. L. Grotch, ibid., 43, 1362 (1971).

7a. Applications of Artificial Intelligence for Chemical Inference. I. The
    Number of Possible Organic Compounds: Acyclic Structures Containing C,
    H, O and N.
        J. Am. Chem. Soc., 91, 2973 (1969)
        By J. Lederberg, G. L. Sutherland, B. G. Buchanan, E. A. Feigenbaum,
        A. V. Robertson, A. M. Duffield and C. Djerassi.

7b. Applications of Artificial Intelligence for Chemical Inference. II.
    Interpretation of Low Resolution Mas Spectra of Ketones
        J. Am. Chem. Soc., 91, 2977 (1969)
        By A. M. Duffield, A.V. Robertson, C. Djerassi, B. G. Buchanan,
        G. L. Sutherland, E. A. Feigenbaum and J. Lederberg

7c. Applications of ARtificial Intelligence for Chemical Inference. III. Al-
    iphatic Ethers Diagnosed by Their Low Resolution Mass Spectra and NMR
    Data.
        J. Am. Chem. Soc., 91, 7440 (1969)
        By G. Schroll, A. M. Duffield, C. Djerassi, B. G. Buchanan,
        G. L. Sutherland, E. A. Feigenbaum and J. Lederberg

7d. Applications of Artificial Intelligence for Chemical Inference. IV.
    Saturated Amines Diagnosed by Their Low Resolution Mass Spectra and
    Nuclear Magnetic Resonance Spectra.
        J. Am. Chem. Soc., 92, 6831 (1970)
        By A. Buchs, A. M. Duffield, G. Schroll, C. Djerassi, A. B. Delfino,
        B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum and J. Lederberg

7e. Applications of Artificial Intelligence for Chemical Inference. V. An
    Approach to the Computer Generation of Cyclic Structures. Differentiation
    between all the Possible Isomeric Ketones of Composition $C_6H_{10}O$.

Org. Mass Spectr., 4, 493 (1970)
By Y. M. Sheikh, A. Buchs, A. B. Delfino, G. Schroll, A. M. Duffield,
C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum and
J. Lederberg

7f.   Applications of Artificial Intelligence for Chemical Inference. VI. An
      Approach to a General Method of Interpreting Low Resolution Mass Spectra
      with a Computer.
      Helv. Chim. Acta., 53, 1394 (1970)
      By A. Buchs, A. B. Delfino, A. M. Duffield, C. Djerassi, B. G. Buch-
      anan, E. A. Feigenbaum and J. Lederberg

7g.   The Application of Artificial Intelligence in the Interpretation of Low
      Resolution Mass Spectra.
      Advances in Mass Spectrometry, 5, 314, (1970)
      By A. Buchs, A. B. Delfino, C. Djerassi, A. M. Duffield, B. G. Buch-
      anan, E. A. Feigenbaum, J. Lederberg, G. Schroll and G. L. Sutherland.

7h.   Applications of Artificial Intelligence forChemical Inference. VIII.
      An Approach to the Computer Interpretation of the High Resolution Mass
      Spectra of Complex Molecules. Structure Elucidation of Estrogenic Steroids.
      J. Amer. Chem. Soc.,
      By D. H. Smith, B. G. Buchanan, R. S. Englemore, A. M. Duffield,
      A. Yeo, E. A. Feigenbaum, J. Lederberg and C. Djerassi

7i.   An Application of Artificial Intelligence to the Interpretation of
      Mass Spectrometry.
      By B. G. Buchanan, A. M. Duffield and A. V. Robertson,
      Mass Spectrometry, B. W. G. Milne, Editor, John Wiley and Sons,
      New York, 1971. pp. 121-178.

8.   D. H. Smith, A. M. Duffield - d C. Djerassi, Org. Mass Spectrom.,
     Submitted for publication.

9.   Anal. Chem., 42, 1122 (1970); W. E. Reynolds, V. A. Bacon, J. C. Bridges,
     T. C. Cobum, B. Halpern, J. Lederberg, E. Levinthal, E. C. Steed,
     and R. B. Tucker.

(4) <u>Cell Separator Project</u>.   Prepared by L. A. Herzenberg and E. Levinthal.

The Cell Separator Project, currently in its third year, is developing the equipment and techniques for automated high speed sorting of functionally different human and other mammalian cells. This project involves an interdisciplinary group of biologists under the direction of Dr. Leonard Herzenberg, Professor of Genetics as well as a staff of engineers and supporting technicians located in the Instrumentation Research Laboratory under the direction of Dr. Elliott Levinthal, Senior Research Scientist. The biomedical objectives include:

(a)   Separation of the various cells involved in the humoral (antibody) and cell mediated (hypersensitivity) immune response. Antigen binding cells, thymus derived cells, bone-marrow derived cells, cells with specific immunoglobulins on their surface will be detected and viably separated after appropriate immunofluorescent surface staining.

(b)   Study the binding kinetics and affinities of cell surface molecular probes like Concanavalin A, other phytoagglutinins, and aniline napthalene sulfonate (ANS) with the aim of distinguishing and separating different cell types including perhaps malignant from normal cells.

(c)   Select somatic cell intra or interspecific hybrids after Sendai virus fusion by nondestructive positive immunoselection.

(d)   Detection of fetal red blood cells in maternal circulation.

(e)   Differentiating leucocytes and other cell types in normal and pathological body fluids.

(f)   Detection of tumors by reaction of circulating tumor cells with fluorescent labelled tumor specific antigens.

(g)   Other related applications on an opportunistic basis, as it becomes apparent that such work is worthwhile.

The instrumentation effort involves the development of the optical flow system and separator components as well as the control electronics and software, for the cell separator. The instrument consists of a nozzle assembly designed to provide examination of single particles flowing in a narrow stream and a pulsing and deflection assembly designed to physically separate particles of interest from other constituents of the stream.
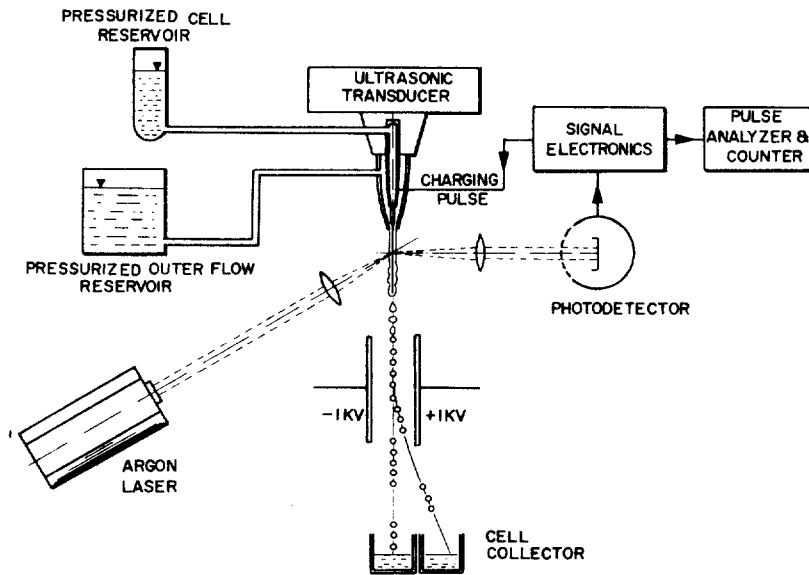
Figure Cell Separator-1

Simplified Block Diagram of
Cell Sorter

(a) Background

In the same way that many of the spectacular advances in molecular biology
were impossible until it became possible to separate functionally different
molecules by such techniques as electrophoresis and ultracentrifugation,
advances in cell biology have awaited development of instrumentation
able to separate large numbers of functionally different cell types.
Many have attempted to do this by bulk methods, but the resolution of such
methods is limited. It appeared to us that the best approach to the
problem was to inspect the cells individually and sort them on the basis
of these individually measured characteristics. We have found that a
number of separations of biomedical interest could be accomplished using
fluorescent markers on the desired cells and electronically deflecting
drops containing the various types of cells into separate containers.
The only other workers with a similar approach are Fulwyler, et al., who
have demonstrated electronic cell sorters operating on volume (1) and are
now building a unit able to operate on both volume and fluorescence
(personal communication). Several workers have described cell analysis
systems based on flow techniques similar to those used in our equipment
(Van Dilla, et al. (2), and Kamentsky et al. (3).

Biophysics, Inc., ("Cytograph" and "Cytofluorograph") and Technicon
Instruments Corp. ("Hemolab D") now market cell analysis instruments
using flow techniques but these instruments do not have separation capability.

(1)   Fulwyler, M. J., Glascock, R. B., and Hiebert, R. D. "Device
       Which Separates Minute Particles According to Electronically
       Sensed Volume". Rev. Sci. Inst. 40: 42, 1969.

(2)   Van Dilla, M. A., Trujillo, T. T., Mullaney, P. F. and Coulter,
       J. R. "Cell Microfluorometry: A Method for Rapid Fluorescence
       Measurement". Science 163: 1213, 1969.

(3)   Kamentsky, L. A., Melamed, M. R., and Derman, H. "Spectropho-
       tometer: New Instrument for Ultrarapid Cell Analysis". Science
       150: 630 (1965).

Since starting the major effort on this project in 1968 a number of
significant successes have been achieved in both the technical development
as well as the biological applications. The following bibliography
provides a summary of these results:

(1)   W. A. Bonner, H. R. Hulett, and L. A. Herzenberg. "Highspeed
       Sorting of Fluorescence Labeled Cells", Fed. Proc. 30, 699
       Abs. 1971.

(2)   L. A. Herzenberg. Chairman, Conference Session, "Fluid
       Transport Methods", Engineering Foundation Research Conference
       on Automatic Cytology, New England College, Henniker, New
       Hampshire, July 26-30, 1971.

(3) L. A. Herzenberg and R. G. Sweet. "Fluorescence Activated
     Cell Sorting", presented at Engineering Foundation Research
     Conference on Automatic Cytology, New England College,
     Henniker, New Hampshire, July 26-30, 1971.

(4) L. A. Herzenberg, T. Masuda, and M. Julius. Invited paper
     on Symposium on Thymus and Bone Marrow Cells in the Immune
     Response, Annual Meeting of the American Society for
     Hematology, San Francisco, Dec. 4, 1971.

(5) L. A. Herzenberg. Invited participant in symposium on Cell
     Purification by Use of Surface Antigens and Receptors,
     Midwinter Conference of Immunologists, Asilomar, California,
     Jan. 22, 1972.

(6) L. A. Herzenberg, R. G. Sweet, M. Julius, T. Masuda, and
     R. A. Merker. Invited paper, "Fluorescent Activated
     Electronic Cell Sorting in Immunology", to be presented at
     Biophysical Society Annual Meeting, Toronto, Canada, Feb.
     19, 1972.

(7) W. A. Bonner, H. R. Hulett, R. G. Sweet, and L. A. Herzenberg.
     "Fluorescence Activated Cell Sorting", Rev. Sci. Inst. 43,
     404, 1972.

(8) L. A. Herzenberg in "Immunological Intervention", Jonathan
     Uhr and Maurice Landy, eds. Academic Press. (In press, 1971).

(9) M. Julius, T. Masuda and L. A. Herzenberg. "Isolation of Functional
     Antibody Forming Cell Precursors Using a Fluorescence Activated
     Cell Sorter". (In preparation).

(b) Rationale

The rationale behind our approach was simply that separation of large
numbers of functionally different cells would make it possible to conduct
many important studies on specific cell functions. In order to acquire
large numbers of cells in a reasonable time, rapid observation was
necessary. This effectively eliminated scanning systems and limited us
to use of only a few parameters. A flow system was a logical way to
look at the cells rapidly and sequentially. Use of fluorescent techniques
provided readily available means of differentiating between many
funtionally different types of cells, but required incorporation of a
laser light source in order to provide sufficient signal-to-noise ratio
to detect the cells. Electronic sorting techniques originated by Sweet
and adapted by Fulwyler (see Background) provided us with a basis for
developing a rapid, accurate method of sorting desired cell types as a
function of fluorescent information.

(c) Methods and Procedures

The procedure involved in the use of the cell separator presently involves three steps:

A. Preparation of cells occurs over a period of hours or days depending on the experiment. Cells of interest, immunologically sensitive cells for example, are collected and tagged with a fluorescent marker.

B. This single cell suspension is then brought to the instrument and analyzed and/or divided into fractions. This latter procedure involves a certain amount of subjective decision making by the experimentor and machine operator. Data on the cells, e.g. distribution of several fluorescent and low and wide angle light scattering amplitudes is acquired before separation for analysis and then thresholds and/or windows for the various parameters are set for separation. Data on the cells, usually the distribution of their fluorescent signal amplitudes, is acquired before or during separation. The resulting fractions are sometimes reexamined via the cell separator or by microscope. More frequently the fractions are tested by the Jerne plaqueing technique or reinjection into irradiated hosts.

C. Finally, the data collected from the cell separator (stored on the computer) and resulting biological procedures are correlated statistically.

The scientific gains to be made by collaboration with SUMEX are both improved operations under B by on-line coupling of the instrument to a computer to allow interactive decisions to be made during separations and under C by more sophisticated statistical analysis of the data.

(d) Significance

Separations of functionally different, viable cells permit their characterization and studies of their function and interactions. It is as crucial a step in cell biology as precise methods for protein and nucleic acid separations in molecular biology. Progress in understanding the immune system is being speeded by separation of functional cells of the lymphoid system.

Development of rapid machine assisted hematological, and cell pathological diagnostic methods will increase clinical laboratory capability and decrease the cost of current manual methods.
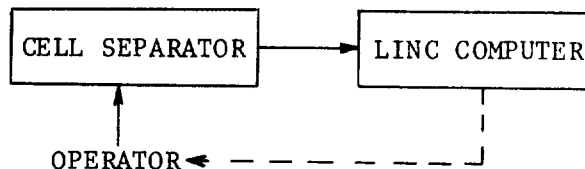
(e) Computer Interaction

As mentioned, in the background section, there are only one or two other groups successfully pursuing this approach to cell characterization and separation. This is due to the requirement for success of a juxtaposition of skilled biological and engineering personnel. Stanford is also especially fortunate in having a very active program in computer development.

The cell separator project is currently using a dedicated small computer on-line (LINC) as well as the ACME system on a less regular, off-line basis.
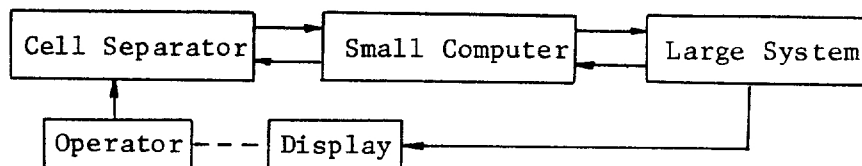
There are two levels of interaction between the cell separator(s) and general purpose computers.

A. The on-line, hook-up of each instrument to a small computer provides a data collection and analyzer system for preliminary results. This is currently in use full time. It allows 4 to 8 hour experiments with assurance that the data is meaningful at each step of the experiment. The human operator is presently the primary link in the feedback loop from the data collection system and the experimental equipment.

```
┌──────────────────┐        ┌──────────────────┐
│  CELL SEPARATOR  │───────▶│  LINC COMPUTER   │
└──────────────────┘        └──────────────────┘
            ▲                          │
            │                          │
      OPERATOR◀ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┘
```

Such interaction requires feedback times in the range of a few minutes. More rapid feedback between the small computer and cell separator is envisioned as the characteristics and uses of the instrument become better understood.

B. The second level of interaction, based on the capabilities of a larger computer system, will provide more sophisticated analysis of results.

C. A third level would use the larger computer to enhance the functioning and software developments necessary for the optimimum use of the small computer.

```
┌───────────────┐     ┌───────────────┐     ┌───────────────┐
│Cell Separator │────▶│Small Computer │────▶│ Large System  │
└───────────────┘     └───────────────┘     └───────────────┘
        ▲                                            │
┌───────────┐      ┌───────────┐                     │
│ Operator  │─ ─ ─ │  Display  │◀────────────────────┘
└───────────┘      └───────────┘
```

Multiparameter analysis of cell distributions is one of the features which would be of immediate benefit to the project.

Implicit in our need for multiparameter analysis, is a feedback time scale which allows modification of experimental procedures. For example, a three-dimensional display of 2 or 3 minutes of data offers the opportunity of more precise manual adjustment of separation parameters. Also, a numericalintegration of areas under three-dimensional curves could actually be used to automatically, and continuously readjust these separation controls. The result of such feedback will be more consistent and more precise definition of the cell populations separated and studied. The large system's language capability can dramatically increase the feasibility of "programming" the instrument to carry out various experimental regimes. At present, the effort involved in programming the small computer has forced the project to use a standard set of software routines and tailor the experiment to fit these. Increasing the coupling between the small computer and cell separator has received limited emphasis because of the software developments required.

After some years of experience with realtime use of computing facilities, we feel that the above uses are not only possible, but feasible without our taking on the role of a major computer development project ourselves.

(5) <u>Average Evoked Potentials and Perception - prepared by Bert S. Kopell and Walton T. Roth</u>

(a) <u>Problem Statement</u>

The Laboratory of Psychophysiology at the Palo Alto VAH has been studying human neurophysiology as related to psychiatric disturbance and the action of medications for about eight years. We have even been able to measure drug effects neurophysiologically at doses below the threshold for subjective effects in the cases of cortisol (Kopell <u>et al</u>., 1970a) and thyroid hormone (Kopell <u>et al</u>., 1970b). A major emphasis has been on finding neurophysiological measures of perceptual processes. Very little is known about the realtionship between subjective perception and its neurophysiological counterpart, but there is some indication that the average cortical evoked response afford an objective index of perceptual process and attention (Satterfield, 1965). Various perceptual processes are known to be disturbed or altered in psychiatric disorders and by psychoactive drugs. Objective neurophysiological measurements of subjective perceptual processes may have the potential for giving prognostic indicators regarding psychiatric decompensation or the predilection for alcohol or drug addiction.

(b) <u>Background and Rationale</u>

Our primary interest is in the electrocortical activity in response to sensory stimulation. The averaged cortical evoked response has been shown to be affected by various physiological and psychological parameters including medications, alcohol, attention, and psychiatric disturbances. A small electrocortical response is generated with any sensory stimulation but this cannot be seen in the EEG as measured directly from the scalp. Computer averaging of the EEG over several stimulations, however, can be utilized to enhance the size of the signal and make it available for study. Recently various members of our laboratory have used the classic LINC and the PDP-12 for this purpose (Gips <u>et al</u>., 1972). While these computers are an advance on the technology of 10 years ago, they do have limitations. Just as these computers have provided for more sophisticated experimentation as compared to the special purpose computers, the proposed system will allow us to make a qualitative and quantitative advance in the complexity and sophistication of our investigations.

A particular problem in studying perception in this manner is presented by the fact that perception is an evanescent and continuous process and the evoked response is the stochastic result of an accumulation of this process over time. Very rapid on-line statistical manipulations of the EEG are needed if one is going to achieve a better understanding of the perceptual process utilizing this technique. With this proposed computation power, one can obtain a closer estimate of the perceptual process from second to second rather than only being able to talk about the "average" process over time. Indeed, further knowledge of the nature of the second to second variations is crucial in understanding the perceptual process.

(c) <u>Methods and Procedures</u>

The data acquisition methods that are needed in our experiments are in the order of 10 KHz and above sampling rate over a period of up to an hour. Though the PDP-12 can acquire data at this rate, it cannot either store or perform sufficient analytic operations on data at this rate. A solution to this problem is to use the PDP-12 as a peripheral to a larger computer (PDP-10) and perform the sampling and analog to digital conversion with the PDP-12 and the analysis and storage with the PDP-10.

An example would be the recording of the EEG from several locations on the scalp of a subject while he is viewing a given set of stimuli. By comparing successive responses to the cumulative response of previous experiments or a given segment of the current experiment, one can determine if he has been able to discriminate a change in stimulus intensity or if his attentional state with regard to the stimuli is changing. The comparison might require doing a statistical procedure such as the Pearson Product Moment Correlation or Fourier Analysis between the digitized EEG for the last second recorded and a previous similar sample for each of the six or eight electrode placements being used. The results of this on-line statistical analysis might then be used to determine certain qualities of the next stimulus to be presented. In this way we can have an on-line feedback controlled perceptual experiment based on a statistical analysis of the EEG. This is not possible with the PDP-12 alone because of the time and data storage requirements.

The major need for the PDP-10/12 configuration is to provide for adequate computer ability to perform on-line analytic procedures of high rate physiological samples and then use these results to alter the stimulus properties. This computer configuration will also allow us to produce complex visual stimuli (such as animated figures) on a computer driven cathode ray tube and simultaneously measure the neurophysiological responses.

The configuration needed includes a PDP-12 which we have from other funding sources. All interfacing with the PDP-10 must be provided for by this grant. This is to include an I/O capability of 10 KHz 12 bits wide (10,000 12 bit PDP-12 words per second) probably implemented via the PDP-12 accumulator and a single I/O command. In addition, a parallel remote graphic display terminal with CRT, keyboard, and hardcopy, which can be used to access the PDP-10 independent of the PDP-12 at least 110 baud is needed. Since our laboratory is physically about two miles from the proposed location of the PDP-10, adequate and reliable transmission facilities must also be provided.

These facilities will include 40K baud telephone lines and a remote PDP-11/05 data concentrator. If a 4:1 data compression is possible (a likely case), a single telephone line with synchronous modems will suffice. Otherwise one, up to an unlikely three, additional units will be needed. The CRT keyboard would be standard.

(d) Significance

The experiments that are being performed in our laboratory are designed to increase our knowledge of the action of drugs such as marihuana, alcohol, heroin, and methadone. Electrophysiological data correlates with central nervous system processes that are independent of language and other social response variables. Thus, such data can be generalized to drug users from different backgrounds. In addition, information obtained by physiological methods is appreciated by the general public as being objective and trustworthy, and can be used to reduce the credibility gaps between the scientific community and drug users.

In the case of marihuana little is known about its effects on attentional processes, which are of obvious relevance to such activities as driving an automobile. Many users claim that they can control their mental state while intoxicated with marihuana and perform normally if need be. With neurophysiologic measures of attention the claims can be objectively tested.

Addictive drugs such as heroin or alcohol raise other questions. For example, why do some persons become heavily dependent after experimenting with these drugs, while others do not, despite equivalent experimentation? Obviously socioeconomic and environmental factors play a role, but there also may be physiological differences in susceptibility. Another question is what is the most useful treatment program for a given individual? The individual differences we are looking for may allow rational decisions to be made as to whether a heroin addict should be put on a program of abstinence or methadone replacement, or whether neither of these alternatives is likely to succeed. Since there is a center for the treatment of heroin addicts here at Palo Alto Veterans Administration Hospital, we have a readily available source of experimental subjects that can be followed long enough to test prognostic predictions. Another ward specializes in the treatment of alcoholics, and it has provided a place for us to study an experimental cycle of intoxication and withdrawal under controlled conditions. In the case of alcoholics there are many pressing questions that may be approached neurophysiologically. As in the case of heroin addicts, there is the problem of differences of susceptibility with its implication for prognosis and the choice of treatment. Also there is the question of whether occult brain damage from chronic abuse of alcohol is present in a given patient. The determination of an alcoholic deficit has important medico-legal implications, as well as an implication as to how much rehabilitation is possible.

Thus, the investigations that we are undertaking are very important in view of the expanding use of drugs of all kinds. Aside from the obvious theoretical yield from studying the neurophysiological actions of psychoactive compounds, there are immediate social and treatment implications.

(e) Relationship to Other Work

Our work involving alcohol and heroin addicts is being performed in conjunction with Dr. Zarcone and Dr. Dement, members of the faculty of the Department of Psychiatry. They are especially interested in mechanisms of sleep dysfunction and hallucinosis in alcoholics in relation to serotonin metabolism. Our marihuana work has enjoyed the collaboration of Dr. Tinklenberg, also of the Department of Psychiatry, who is interested in measuring psychological changes simultaneously with neurophysiological ones. He has developed some special techniques for measuring memory and attention in drug states.

There are very few centers in the world which do work of a nature similar to ours. Dr. M. Buchsbaum at NIMH has been investigating evoked response correlates of perceptual process for several years, and has provided us with some valuable techniques. Dr. C. Shagass of the University of Pennsylvania Medical School has been a leader in studying aspects of the evoked responses that are relatively independent of attention and memory in patients with mental disease and patients under the influence of drugs. The design of some of our studies on alcoholism is based on work being done at St. Elizabeth's Hospital in Washington, D.C., by Dr. N. Mello under the direction of Dr. Morris Chavetz.