TABLE OF CONTENTS

VIII. APPENDICES

NOTE

This last annual report, covering FY73, ACME's administrative extension year, was prepared at this time (May 1973) because staff was available to do the work and because NIH requested that the report be made.

An addendum to this last annual report will be prepared sixty to ninety days after the end of the grant period (July 31, 1973).

NATIONAL INSTITUTES OF HEALTH DIVISION OF RESEARCH RESOURCES BIOTECHNOLOGY RESOURCES BRANCH RESOURCE IDENTIFICATION

Report Period: Grant No.

From: August 1, 1972 To: July 31, 1973 RR - 311 - 06S1

Date of Report Preparation:

May 1973

Name of Institution's Biotechnology Resource Advisory Committee:

ACME Policy Committee

Membership of Biotechnology Resource Advisory Committee:

I. SUMMARY

A. Brief Evaluation.

The ACME facility is now approaching the end of a six-year grant plus a one-year administrative extension. The ACME experiment has proven highly successful as demonstrated by the following points:

1. Teachability:

Medical researchers have been taught to do their own programming for non-trivial tasks. More than 1700 persons on the Stanford medical scene have been trained in the use of computers.

2. Strong Educational Tool:

This facet of ACME encourages many persons to become involved in computing.

3. Data Acquisition:

The ACME system combines moderate rate data acquisition service with timesharing. A relatively sophisticated group of realtime data acquisition users has been developed.

4. User Community:

More than 210 user projects, exclusive of ACME staff, are current users of the system. They enter the system from 55 terminals spread throughout the Medical Center.

5. Programming Effort:

The ACME disk packs and tapes hold programs representing over 250 man years of programming effort.

6. Publications:

A list of recent publications by ACME users and an index of ACME Notes prepared by ACME staff are presented in Section VIII.

7. Dedicated Systems:

Several groups are now using dedicated computer systems which reflect an outgrowth of pilot projects performed on the ACME facility. We observe a propensity of large or clinical projects to become autonomous from large central facilities.

The ACME experiment was initiated nearly seven years ago to provide a timesharing system on an IBM 360 hardware system concurrent with some realtime data handling. Applying hindsight to the choices made we can now see that our successes noted above are mitigated by some trends which were not predicted by us in 1965. Specifically, technological and price/ performance changes have occurred since 1965 which make mini-computing systems dedicated to specific tasks much more competitive with centralized resources than was true in 1965. Based upon our experiences, we at Stanford Medical Center are moving to establish an improved large central computing system while, at the same time, the number of dedicated mini-systems is steadily increasing. In other words, the clear advantages of a large central system for certain applications are offset by other goals in a number of realtime and control situations where dedicated mini-systems offer the only realistic solution.

Some of the more significant changes over the past seven years have included the following:

- (1) Both logic and, somewhat later, main memory units have dropped in cost by 1 to 2 orders of magnitude for devices of roughly equal performance.
- (2) Disk systems have increased in capacity, speed of access, and reliability while costs have dropped markedly. The figures below demonstrate this trend.

DISK PERFORMANCE DATA FOR IBM HARDWARE*

* Data received in IBM presentation in San Jose, California held in May, 1973.

- (3) The prices for mini-computers have dropped an order of magnitude while capability and level of software support have increased significantly. Also, the variety and flexibility of adding various peripherial devices has risen sharply.
- (4) Communications terminals and supporting equipment has improved in terms of speed and reliability.
- (5) The user community has grown in numbers and level of sophistication. It also demands increased availability and reliability. Response times must now be measured in fractions of seconds rather than seconds which imposes high overhead costs on large time shared systems.

In 1965 a handful of computer users existed in the Medical Center. Today, there are more than 200 active research projects on the ACME system plus about 25 mini-systems in laboratories. Computers have become accepted for production use in many applications.

The first year and a half of the ACME grant was spent assembling staff and hardware and developing the PL/ACME system. The result was one of the first timesharing systems with concurrent realtime support. The system we mounted is remarkedly easy to learn and use. Since the system has not been exported, we assume that a shift back to the mainstream of software systems as provided by vendors will become necessary. The relative cost of people versus hardware has grown to a point where " home brew" systems cannot be afforded over the long term. This situation is regretted since many of the newly announced systems fail to deliver to the enduser the convenience and power of our existing system. At Stanford, the conversion to vendor-supplied systems might be expected to occur in three or four years.

The series of developments in mini-systems has relieved some burdens but created new needs for small machine support from a central system. Stanford expects to expand upon the currently available intercommunication systems. Since mini-systems are frequently being used as data collection controllers, the development of shared data base systems on the central machine requires intermachine communication capability.

We have offered graphics services on the ACME system over the past five years. There are now six graphics CRT's plus five hard copy plotters attached to the ACME system. The growth in graphics usage has been slower than one might have expected. We suspect that this reflects a lack of description tools, the relatively high cost of graphics terminals, plus the high cost of running graphics software. Perhaps, the growth spurt will come in a few years after

the costs are dropped another factor of 5 or so, and description tools are improved.

Our dependency upon NIH in establishing a user community and developing a financial base with which to support computing services warrants special mention. It is clear to me that the venture capital to establish a system, train users, and form a critical mass of support could not have been raised by incremental growth through charges to individual research grants. A central facility development grant, such as ACME grant, is the only available counterpart to venture capital in a cash-accounted grants system. This investment has now paid off: We have a common language in use by nearly every department in the Medical School, a cadre of trained programmers, and a strong momentum in the direction of shared data bases and shared programs. Some continuing incremental development effort will be needed to prevent atrophy.

The passage of time has brought about needs for realtime support systems which exceed the capabilities of our local ACME system by one or two orders of magnitude. Hather than attempt to build realtime systems which IBM hardware has not been designed to perform, we prefer to rely on alternative systems specifically designed for this use and build improved communications into the central site. Further efforts will therefore constitute another generation of system planning, based on vendor-furnished modules.

We have all heard of TSS, TSC, TORTOS, CPS, CMS, and other timesharing systems built on IBM hardware systems. It is a credit to the small staff which built PL/ACME that their system can compare so favorably with the other systems which have clearly had far more effort spent on development. The use of PL/ACME as a research tool by so many local groups is a tribute to the system designers and implementers. Gio Wiederhold deserves special credit as the principal creator of the system. He would also be the first to point out some of the design features that might have been improved with the benefit of hindsight. The lists of publications and technical notes appended to this report attest to the productivity of the ACME Facility staff and to the effectiveness of the tools provided to the user community.

J. Lederberg

B. Highlights of FY73,

1. Planning for the Future.

During the past twelve months the ACME computing facility has passed through an identity and existence crisis. The sizeable effort expended by faculty, Hospital staff, and ACME facility staff has led to a decision to maintain the PL/ACME system on new hardware. The computer services for Hospital administration and ACME time sharing and realtime data acquisition services are being merged onto a new facility to be installed in August 1973. Numerous studies and presentations have been required to bring about the decisions which make this possible. This subject is discussed in detail in Part C of this section and in Section II.

2. A Generalized Time Oriented Database System.

Special attention should be focused on the transition which has begun to occur at Stanford with respect to faculty attitude toward the need for sharing of data. The awareness of shared database concepts has increased markedly. Evidence of this can be seen in the teamwork demonstrated in preparation of a health care resources research proposal. Other evidence can be seen in the attendance at seminars concerning the Time Oriented Database system (TOD), developed this year by the ACME staff in conjunction with Dr. James Fries of the Division of Immunology. For more on this subject, see Section V.

3. Software.

System programming development activities during the year resulted in new data compression routines, file system improvements, mounting the COBOL compiler, studies and planning concerning the new VS2 system announced by IBM, support for the small machine multiplexor, and a PDP-11 simulator.

4. Hardware.

The small machine multiplexor was completed, allowing for inter-machine communications. Other hardware projects included work on terminal lightboxes and several new interfaces for users of the 1800 and the multiplexor. ACME acquired several 300 and 1200 baud terminals during the year. In April 1972, we installed a Memorex terminal controller which has performed very well.

5. Core Research.

Support of core research and development effort included programming and computer service support for the DENDRAL project, assistance for the Drug Interaction project, direct support of the initial application of the Time Oriented Data (TOD) system, extension of small machine support to GC/MS activity, a joint development effort on communication hardware development, and a core project to develop new statistical analysis techniques.

6. Utilization.

Utilization of the ACME system in terms of terminal hours has remained relatively constant during the past year. One exception to this is the Drug Interaction project in the Pharmacy which used ACME extensively through February 1973, after which time the system was moved to a dedicated dual mini-computer system.

Since the follow-on to the ACME system was not resolved until March 1973, the rate of new user signups has dropped from normal levels and there have been essentially no new realtime users of the system. This is most understandable since many users felt that ACME might not survive beyond the end of the grant period. It is noteworthy that the user community has continued to use the system in the absence of (prior to March 1973) any Medical Center commitment to retain the PL/ACME system beyond July. Now that such assurance has been made, new users are again expressing interest, evidenced by the number of signups for the introductory classes in use of PL/ACME.

7. Minicomputers.

Other computing activity in the Medical Center includes the acquisition of several minicomputer systems for various research and production projects. Approximately thirty minicomputer systems are currently used within the Medical Center. Some of the applications include data acquisition for mass spectrometers, operation of the Drug Interaction programs, an information system for the Clinical Laboratories, and research support in Nuclear Medicine, Chemistry, Psychiatry, Cardiovascular Surgery, Cardiology, and other divisions and departments. The growth and number of minicomputer systems used for instrumentation control and data collection have pushed the central facility to provide small machine communications and other support activities.

8. Documentation & Conversion.

Throughout Fiscal 73, the staff has spent a great deal of time on documentation of the existing ACME system. Since the decision to move to a merged/l58 facility, the conversion effort has been of central importance.

C. Planning and Reorganization.

1. Stanford Center for Information Processing (SCIP).

In the past Stanford operated five major service computing organizations, each of which had its own loyalties to a specific user community. The five were:

Stanford Linear Accelerator Center University Administrative Computing Facility Campus Facility ACME Facility Hospital Data Processing Facility

Computing at Stanford University was reorganized during the spring of 1973. The new organization, entitled "Stanford Center for Information Processing (SCIP)", provides a unified structure for the five facilities mentioned above. ACME and the Hospital Data Processing Facility will be combined to form the Medical Center Computing Service (MCCS). The heads of all the facilities will report to the director of SCIP. Along with the reorganization of staff involved in managing the various computer facilities, the policy committee structure comprised of faculty members is currently being modified.

2. The Medical Center Planning Effort.

A description of the computer planning activity at the Stanford Medical Center over the past eighteen months would fill many volumes. Several different faculty committees and staff groups have reviewed alternatives ranging from highly distributed interconnected minicomputer systems to highly centralized large computing systems. The issues faced by the various groups were:

- 1. Should PL/ACME service be continued?
- 2. Can the ACME users provide a critical mass of dollars required for a stand-alone facility.
- 3. If a merger is required, who should be the parties to the merger?
- 4. Are the potential advantages of a shared database between Hospital and Medical School strong enough to outweigh the disadvantages of merging a production system with a research support system?
- 5. Should realtime computing services continue to be provided from a central computing source?
- 6. How should computing at Stanford University and the Medical Center in particular be organized?
- 7. What computing services will be needed over the next several years?
- 8. How can we relate Medical Center computing planning to broad University goals?

3. The 370/158.

These are among the many issues which have been considered during the past year and a half. The solution selected entails the installation of an IBM 370/158 hardware system using IBM's newly announced VS2 software system. The services to be offered will include batch services in several languages, time-sharing using PL/ACME, realtime data acquisition services using the existing 1800 system, normal consulting and user services, and small machine communications.

The current schedule calls for removal of the 360/50 system from Stanford on July 28, 1973. A number of peripherals will be moved to the new 370/158 site where systems programmers will have approximately 2-l/2 to 4 weeks to bring up the new system. We expect to resume PL/ACME services for terminal users by September 1, 1973; realtime services will hopefully be available approximately one month later.

Funding for use of computers within the Stanford Medical Center is expected to drop over the next eighteen months due to cuts in federal budgets as well as escalation of costs within fixed budgets. A tight dollar economy coupled with multiple options for the users (e.g. outside time sharing service, Campus computing facility, more powerful dedicated mini systems) will force the new Medical Center Computing Service to perform very well to attract the business of the Medical Center commuity.

4. The SUMEX Proposal.

A proposal has been submitted by Dr. Lederberg calling for the formation of a Stanford University Medical Experimental Computing Facility (SUMEX). If approved, this proposal would result in the acquisition of a PDP-10 to support a national facility specializing in tools for the development of artificial intelligence in medicine (AIM). The ACME experience has been invaluable in demonstrating both the opportunities and the problems of community-shared resources. In particular it has given us the technical expertise needed to design realistic specialized instruments to serve geographically dispersed but intellectually convergent users.

D. Overview of ACME Experiment.

The ACME experience indicates that a large central resource can provide a very valuable service for users requiring text editing, numeric calculations, statistical analyses, and realtime data acquisition at relatively low rates. Our experience has also demonstrated that a large central facility should not undertake high data rate realtime data acquisition and closed loop control functions if it intends to service a large number of time sharing users concurrently. In addition we have learned that an extensive amount of "handholding" is needed to serve the research scientists in a medical community. This may change in the future when MD's will routinely receive more training in computer science in the course of their college educations.

ACME's initial proposal included the following paragraph concerning hardware selection and resource allocation:

"The IBM/360-50 has been selected for the initial realization of ACME (1) as a machine technically appropriate to the immediate tasks in mind and (2) for its system compatibility with the 360-67 already selected for the eventual replacement of the 7090 by the Stanford Computation Center. The 360-50 will be installed in ACME May 1966 and will run on three shifts under Operating System/360, subject to review by the policy committee. These will be dedicated respectively:

- (A) A prompt access time-sharing mode perhaps over most of the working day.
- (B) A scheduled, full-use, on line mode to service development work on high data rate and on line control applications, and for similar systems development.
- (C) Job-shop, especially longer runs for which overnight turnaround is acceptable, and which cannot be serviced with comparable effectiveness by SCC."

The following aims were added to the ACME charter at the time of the Renewal Proposal in the Spring of 1969:

1. To improve hardware and software reliability for the benefit of the medical users.

2. To provide small machine assemblers in PL/ACME so that code for small machines can be written from an ACME terminal.

3. To achieve over time a state where income from user charges will match operational costs for the ACME system.

All of the original objectives have been achieved to varying degrees of satisfaction. Of special note is the development of PL/ACME as an interactive time sharing system which can be easily learned and used by medical staff. On the other hand, the realtime support offered is inadequate due to system instabilities and data rate limitations. Access to Campus Facility is inconvenient for ACME users.

In terms of the items added at renewal time, hardware and software reliability have been markedly improved. Small machine assemblers have been added, but the user must write code in the assembly code for whatever satellite he intends to run. At present, assemblers of this type exist for PDP-11, PDP-8, and 1800. The income of the facility has been rising steadily. Economic overlaps with NIH direct support for ACME have blurred the transition to totally non-subsidized use. A major rate increase was initiated in April, 1972. With this change, income over the last 12 months reached roughly $\overline{55\%}$ of direct operating costs (exclusive of development efforts). From the vantage point of hindsight one could well ask whether the selection of the 360/50 hardware and the decision to promote a large central time sharing and data collection resource were appropriate. Given the availability of new third generation hardware and the promises of IBM or expectations of its customers in 1966, the 360/50 hardware selection is defensible. However, the development of low cost, fast, well-supported minicomputers was not anticipated to proceed at the phenomenal pace that it has. This major technological shift has strongly influenced our present thinking for the future of computing in medicine and related research. The role of a large shared resource has by no means been obviated by the minicomputer revolution. We will continue to need powerful facilities beyond the scene of current mini architecture.

II. STANFORD MEDICAL CENTER COMPUTING PLANS

A. The Current Scene.

Between January and April 1973, the following significant events occurred in the computing environment affecting the Medical Center:

1. The University reorganized the service computing management structure to form the Stanford Center for Information Processing (SCIP). The SCIP organization will manage and operate all major service computing functions for the University.

2. The Board of Trustees authorized acquisition of an IBM 370/158 system to service the needs of the Medical Center.

3. Personnel from the ACME Facility and Hospital Data Processing Facility were assigned the task of converting the current systems to the new hardware systems.

4. Users were notified of the changes scheduled to occur between July and December 1973.

5. Planning of new faculty advisory groups for computing throughout Stanford was done.

PL/ACME users had been warned that the time sharing service might have to be disbanded at the close of the ACME grant. Therefore, the ACME community was elated by the above series of decisions. Medical School faculty and Hospital management were notified of the scheduled changes by a memorandum from Mr. Victor Barber dated April 23, 1973. A copy of this memorandum has been reproduced on the pages which follow.

B. Shared Database Planning.

The need for planning of shared database effort is presented in Appendix A. Given the need to which these memoranda attest, it is likely that the central facility will assign key personnel to work on the problem along with interested researchers. The near term development effort is likely to be based on use of the Time Oriented Database (TOD) system. Further information on TOD is presented in Section V of this report.

C. New Faculty Appointment.

A selection committee has nearly completed its deliberations with respect to a new faculty member in the Medical School who will have considerable responsibility for policies affecting computer services. The new position will be located in the Department of Community and Preventive Medicine. It is hoped that the new appointee will serve as a focus and spearhead for development activity in the shared database area.

- 11 -

DATE: April 23, 1973

 $\overline{1}$ O Distribution

U. A. Bailer

From V. H. Barber, Associate Director, Medical Center Computing Service Stanford Center for Information Processing

SUBJECT Computing Services for Medical Center

Stanford University announced in March of 1973 the complete reorganization of its general support computing facilities. The new organization, Stanford Center for Information Processing (SCIP), is described in the attached press release. The result of the reorganization provides a large general support computing facility in the Medical Center environment that merges the services of the medical research community with the business, administrative, and patient care activities.

My role as Associate Director of the Medical Center Computing Service is to serve Medical Center users and represent their interests in the service computing arena. Our goal will be to provide the required computing services at the lowest possible cost. A new hardware system will be available to serve the entire Medical Center in September 1973. The new facility will have more than three times the compute power of existing service facilities and will make available improved services during the year as the power of the system is harnessed with associated software. These will include shorter response or turn-around time and sharing of data bases.

A list of service goals for the MCCS is attached as Appendix A. Additional needs of the medical community will be established through interaction with the Hospital management group, individual users, and faculty committees.

We intend to provide easy communications between you, the user, and the staff of the new facility. We want a highly personalized service that is responsive to the needs of the medical community. Madhu Bhide,x5151, will be the primary liaison and coordination point for Hospital services, especially those oriented toward financial applications. Ms. Karen Richards, R.N., will continue as the Nursing Service Coordinator for computing matters. She is available at x6084. Ron Jamtgaard, x6121, will be the primary contact for users of timesharing services and realtime support; he will respond to needs of the Medical School research and education functions. B. J. Gaul will be Operations Manager of the 370/158 computing facility. He is available at x5880.

Ron Jamtgaard and his staff will be housed in the old ACME offices $(TC101, temporary building).$ All other MCCS computing and management personnel will be located in the Administrative Services Annex, just north of the Medical Center, in the old Hospital Data Processing area.

We plan to serve you in the following manner: The new hardware facility (IBM 37C/l58) will be available in September 1973 to serve the reasearch and development interests of the Medical Center; by December 1973, the business and finance computing will be merged onto the new system. The services which will be offered when the facility opens include timesharing (using the PL/ACME language) and batch services in those languages for which a user need exists. Initially, batch services will be provided for FORTRAN, COBOL, PL/l, and LISP. Services to the business and finance community will continue as before; however, larger resources will be available, and there will be new opportunities for service.

The transition from PL/ACME on the current Model 50 to the 370/158 will be transparent for the terminal user except possibly in some realtime areas. Our target is to hold service interruption times to a minimum. Standard terminal services will likely be unavailable for about four weeks; realtime services may be disrupted for six to ten weeks. Digital realtime data acquisition services and graphics via the IBM 1800 as well as small machine communications will be provided by the new facility as soon as possible. The transition of Hospital services should be completed by December 1973.

The facility will be operated on income received through user fees. Our goal will be to provide maximum service at the most cost effective rates. Further policy on fees will be developed and released in the near future.

Persons who are new to the Stanford medical computing scene are encouraged to contact me at x5998 so that staff can be assigned to assist in definition and solution of computing needs.

MCCS exists to serve you and your computing needs. We hope to hear from many of you regularly.

- cc: Medical School Faculty Hospital Department Heads ACME Users SCIP Associate Director C. Rich
	- P. Carpenter
	- T. Conda
	- P. Hofmann
	-
	- C. Dickens
	- M. Roberts
	- R. Jamtgaard

MEDICAL CENTER COMPUTING SERVICE FUNCTIONS, SERVICES, AND GOALS

 $1.$ Patient Accounting. Hospital financial and administrative service and patient accounting services. These services involve chiefly patient accounting, patient billing for Stanford Hospital and Clinics, accounts receivable, third-party allocations, payroll, personnel, general ledger, accounts payable, census, financial and budgetary analyses for the Stanford University Hospital.

As part of the patient accounting services, MCCS manages a large patient financial data base. It is expected that this will form a nucleus for a comprehensive patient data base in the future. One of the services that the service facility will offer is building on and managing this patient data base for both the Hospital administrative staff and medical research personnel interested in the patient data base.

- 2. Interactive time-sharing -- PL/ACME. PL/ACME is an easy-to-use time sharing service. The major user of this service is the medical community at Stanford University. It is expected that the proposed facility will continue to support $PL/ACME$ in its present form and gradually enhance the service to satisfy future requirements.
- 3. Data reduction and data control services. Currently a well-trained staff of keyboard operators and data control personnel perform these services for the Stanford University Hospital. This service will continue to be performed, and the recipients of this service will expand from Hospital financial data processing users to include other medical personnel.
- 4 . Realtime services. Realtime service is currently provided by an TRM 1800. and ACME-built interfaces in the laboratories. The 1800 is programmed as an integral part of the system and acts as a 360 control unit. These services are currently used routinely by a number of investigators.
- 5. Small machine communications services. A multiplexor (MUX) for connection of mini-computers to the new $\overline{570}/158$ will be installed in 1973, after more routine services are operational,
- $6.$ Data collection. The new facility hopes to collaborate in the development of data collection systems. An example would be the development of an automated Patient Admission, Discharge, and Transfer system. Opportunities in this field will be vigorously pursued.
- 7. Language support. Other such collaborative efforts are foreseen in the area of language support. One such example is the MUMPS language. This language was developed at the Massachusetts General Hospital. It is used in conjunction with PDP-11 and PDP-15 computers. The Hospital uses a PDP-11 MUMPS system for the Pharmacy Drug Interaction Project. The facility role is not clear; we are open to new ideas here.
- 8. Liaison with Forms Management. MCCS will work closely with the Hospital Forms Management Section. The responsibility of the Forms Management. Section is to coordinate all the forms that are used at the Stanford University Hospital. It takes on the responsibility of designing, printing, ordering, and stocking of forms. This service is expected to continue and will closely interact with the information flow development at the Medical Center.
- 3. i'rogramming and consulting. The proposed facility will offer programming and consulting service. These programming activities include fee-fo zervice programming for users, design and development of production systems. and maintenance of public utility programs as well as existing projuction systems. Additionally, the facility will offer services in the areas of procedures analysis and automation of procedures. One of the current analytical services in which we are participating is an automated work-measurement study. It is expected that the proposed facility will continue to participate in such studies and offer services in these areas.
- 10. Library and grant assistance. Assistance in the identification of fundin opportunities, proposal preparation, and management of grants that include the use of computer facilities will be available.

A library of current reference publications in the area of computer use in the health care field will be maintained.

11. Educational activity. The facility is expected to be very active in continuing its current educational activities, specifically in teaching ACMF's interactive and timesharing usage, as well as education of nursing and other Hospital staff in automated procedures. It is further expected that the proposed facility will extend its activities in continuing education of Hospital staff in data processing procedures, systems design, and data base management, and that it will also be active in the area of current awareness and dissemination of information to physicians and other medical personnel. As an adjunct of this, it is expected that the proposed facility will have an internal awareness program to keep its staff abreast of development in health care technology as applied to a computer service facility of a major medical center.

D. Some Observations on Computer Planning.

It took the Medical Center and University eighteen months to perform the planning activity leading to a decision for a course of action. A chronology of this period is attached as Appendix B. By scanning the chronology one can quickly observe that organizational and technical issues involving computing become quite complex and require extended timeframes to complete. Some of the major policy issues addressed by the various study committees included the following:

1. Do we want a highly centralized computing environment or do we choose a distributive minicomputer system with some inter-machine communications? Would some middle ground between these two choices be most appropriate?

2. Can we successfully merge the research support computing of the Medical School with the business and finance data processing of the Hospital?

3. What advantages might be gained by merging with the central Campus Facility of the University?

4. What investment does the PL/ACME user community have in the PL/ACME system and language? How easily could they be converted?

5. How can we fund computing for medical students and researchers on the faculty?

6. What computing needs are likely to dominate over the next five years?

These are some of the questions which the various committees have addressed.

III. ACME FACILITY ACCOMPLISHMENTS - FY73

Accomplishments of ACME staff personnel are described here; core research projects led by faculty members are included in Section V.

The primary accomplishment of the ACME facility during the past year has been to hold its user community largely intact during a period when the future existence of PL/ACME services was highly in doubt. The doubt stemmed from the fact that PL/ACME services had been subsidized by the ACME NIH grant and that the paying users did not constitute a critical mass to afford a facility which could duplicate these services.

A. Planning Studies.

Since October 1971 several members of the ACME staff have been actively involved in planning methods to continue offering PL/ACME services beyond the period of the ACME grant. There follows a list of some of these studies:

- 1. Merger of Hospital ADP and ACME facilities on a 360/65, PDP-10, or 370/158.
- 2. Merger of University Administrative Computing Facility, Hospital and ACME facilities on a 370/158.
- 3. Merger of Campus Facility and ACME on a 360/67.
- 4. Conversion effort to mount ACME on various systems.
- 5. User surveys to determine user plans if PL/ACME services were dropped.
- 6. Specification of users needs.
- 7. Review of potential need for time oriented database sub-systems.
- 8. Consideration of various organizational alternatives. The results of most of these planning studies have been reported in earlier sections of this report.

B. Time Oriented Database Development.

One of the major tasks of the ACME applications staff during FY73 has been the generalization for ACME users of the Time Oriented Database system originally designed by Dr. James Fries of the Division of Immunology. A lengthy description of this system (TOD) is included in Section V of this report.

C. New and Continuing Applications Programs.

1. DENDRAL:

Support for the DENDRAL project during this fiscal year has consisted of machine services both in interactive PL/ACME and batch LISP. Early in the fiscal year an overnight version of batch LISP was mounted so that jobs could be entered from terminals in the daytime and run when the PL/ACME system was not needed. In addition the LISP interactive compiler was markedly improved. The small machine multiplexor and other small machine support has found limited use in the DENDRAL area. In addition the Loma Linda graphic displays have been fully incorporated into the DENDRAL closed loop control problem.

2. Drug Interaction Project.

This has been the year of transition for the Drug Interaction Project from the ACME system to a dedicated dual PDP-11 system. The new software written in MUMPS is now operational. This project has served as a classic example of how a new idea is formulated by faculty, tested under pilot project status on the ACME system, proposed for a research grant, and finally implemented in a production form. Many of the computing applications in the Stanford Medical Center have followed this course of action.

3. Medical Student Admissions.

Programming is now being done to handle medical student admissions needs. The system will assist the Admissions Office in screening applicants and provide administrative support.

4. Time Series Data Analysis.

Last summer ACME helped to support the work of Dr. Will Gersch of the University of Hawaii who used the ACME system to develop an automatic decision procedure to calculate spectral density estimates. The result of this effort is now available to all users in the form of a public program.

5. Radioimmunoassay Programs.

A number of FORTRAN programs written at NIH have been moved to the PL/ACME system to support research in radioimmunoassays.

6. New Realtime Users.

Very few new realtime users were recruited during the past year. This reflects the doubt in the minds of many concerning the future of PL/ACME realtime services. Two projects which were implemented:

Dr. Don Perkel's project involved analog/digital processing of two to four channels of nerve impulses recorded during swimming of the leech. This is a study of nervous control of movement.

The second project, headed by Dr. P. Sokolove was a study of the role of the nervous system in production and maintenance of circadian rhythms (data consisted of nerve spikes and EXG records).

- Note: Detailed descriptions of Items 1 and 4 above are presented in Section V.
- D. System Software Improvements.

In addition to its considerable planning effort, the systems staff incorporated a number of improvements into the system during 1973. Some of these are listed below:

1. Satellite Machine Support.

The primary effort here involved mounting the small machine multiplexor. This entailed software in the 360/50 as well as small machine code to test the hardware. The new small machine multiplexor can accommodate up to sixteen satellite machines serially passing data to or from ACME (See Appendix C). In addition a PDP-11 simulator was completed. The simulator can be operated in batch or interactively,

2. LISP.

Two tasks were undertaken for LISP users. The first was to mount an overnight LISP batch service to which jobs could be submitted from terminals during the day. The second was to improve the response time of the interactive compiler.

3. File Support.

The primary improvement was a data compression routine which permits users to file their data in a compressed format. For some users, such as those using the time oriented medical record files, this feature permitted a factor of five savings in storage costs. In addition the file system was documented extensively during the last year.

4. Reliability.

A number of bugs were found and fixed. It is a credit to the system staff that we now operate three to four months between system crashes due to software.

5. Other System Tasks.

The number of terminal ports was expanded from 32 to 48 . Accounting programs were modified to capture data at hourly intervals. Batoh accounting was added to the system. Release 20 of OS was implemented. A COBOL compiler was mounted for batch running. The file system directory was rewritten. A terminal survey was conducted to determine the best terminals for ACME to support in the future. A hardware monitor (called the SUM monitor) was attached to the system by Lee Hundley for a series of system measurements. Special programs were provided for ACME-to-OS dataset conversion.

E. Education and Training.

Over the past six and a half years more than 1700 members of the Stanford Medical community have been trained in the use of PL/ACME. This number includes only those who have enrolled in a formal training class. During the past twelve months, nine introductory classes have been offered with a total enrollment of 84. This is less than the normal annual enrollment, primarily due to the fact that classes were not held during three months of the year when the future of ACME was unresolved. Of the 84 persons enrolled, 23% had a Ph.D. or M.D. degree, another 45% had a Bachelors or Masters degree, and 32% indicated no degree. Two years ago the corresponding percentage of Ph.D. and M.D. participants was 35%. Approximately one-half of those who signed up for the introductory class had no prior programming experience; most of the balance had only slight experience. When asked why they sign up for the course roughly 60% indicate their intent to use the computer for numeric calculations and statistical analyses; approximately 20% plan to handle large data files; 10% indicate an interest in realtime applications; the balance require graphics displays and text editing. More than half plan to participate in a project currently using ACME; about 20% intend to start a new project.

It is interesting to note that 100% of the participants report that they have access to an ACME typewriter terminal. There are currently 55 terminals in our network.

In addition to the introductory course, ACME staff have prepared and offered seminars and advanced classes. The seminars have dealt with general medical computing topics. A special series of seminars was held concerning time oriented database work.

Some persons learned to use ACME without enrolling in the formal classes. One might estimate that 30% or more of the current users did this. An aid to the persons who preferred to learn by doing is a new program on the public file called "TEACHER". This is a question-and-answer course designed to be used from an alphanumeric CRT terminal.

F. Hardware Changes.

1. Multiplexor.

The satellite machine multiplexor is the most complex and costly equipment item designed and fabricated by the engineering group during FY73. The device, which can connect up to 16 satellite machines to the ACME system, provides a data path to and from the 360/150 in a demandresponse mode. Data rates of $40,000$ bytes per second are available using 4 twisted pair; rates of 25O,OOO bytes per second are available using coax. The multiplexor is connected to the IEM hardware through a 16-byte parallel data adaptor on tbe 2701 which in turn is connected to a selector channel (See Appendix C). Only three computers have been connected through the multiplexor. Additional customers will not be urged to connect until the new 370/158 system and its software become operational.

2. Terminal Controller.

All terminals are connected to the ACME system via a Memorex 1270 terminal controller which was installed in April 1972. This device has performed very well. It has 32 ports capable of automatic speed recognition up to 1200 baud.

3. Standard Analog Interface Card.

A new standard analog card has been developed for use in the laboratory. Since the 1800 is being moved to a distance 2000 feet further from most users, we will be encouraging the user community to convert to digital signals at the laboratory end rather than in the 1800. As a result, the new standard analog card is not likely to be used extensively.

4. Lightbox.

The standard ACME lightbox which has been used on IBM 2741 typewriter terminals was designed to operate off the 2741 power supply. Our shift to G.E. Terminet 300 terminals and Beehive CRT's has made the lightbox unusable. A new one has been designed and will be placed in use in June 1973.

5. Interfaces for Users.

The engineering group has designed and maintained a number of interfaces for user instrumentation. In general these are paid for directly by the customer. Some of the devices interfaced during the past year include scintillation counters, a paper tape reader-punch, and a Houston plotter.

G. Operations.

The annual average meantime between failures due to all causes (hardware, software, power, and human) reached a new high in FY73: 87.7 hours. A chart presenting additional information on meantime between failures is on the following page.

ACME's Operations Manager, Charles Class, spent a great deal of effort on planning support for the 370/158 facility, providing assistance in the areas of hardware, physical space and communications. Here at ACME he was active in the work of installing the new 300 baud terminals.

Mr. Class was the ACME representative in Co-op, an organization formed by the operations managers of the University's service computing facilities to increase communication and cooperation among the several operations groups.

HARDWARE

These figures do not reflect failures of 1800, PDP-11, or other systems.

Underlined Figures = Best mean time to failure as compared to same period of each year.

+ (May - July, 1973) Projected mean time to failure based upon first nine months, August 1972 through April 1973.

IV. ADMINISTRATIVE ORGANIZATION

The Stanford Center for Information Processing is a new organization for service computing at Stanford University. The SCIP organization is shown schematically below.

STANFORD CENTER FOR INFORMATION PROCESSING

(SCIP).

The ACME computer facility is being merged with the Hospital Data Processing Facility and will now be represented by Mr. Victor Barber as Associate Director of SCIP for Medical Center computing. Until April 1973, the ACME facility was one of three facilities comprising the Computation Center.

The staff of the ACME facility is listed along with the percent of full time equivalent effort on the following page. Major personnel changes which occurred during fiscal year 1973 are as follows:

- 1. Lee Hundley transferred to the SLAC computing facility where he will be working on realtime applications.
- 2. Linda Crouse transferred to the Pharmacology Department as a scientific programmer.
- 3. Rich Cower transferred to the SLAC facility as a computer operator.
- 4. Jane Whitner and Ying Lew were terminated in view of the end of the ACME grant. The University will consider rehiring them as future needs develop.
- 5. Chuck Granieri transferred to SLAC as a systems programmer in the spring of 1973.
- 6. Russell Briggs was assigned full time to the Drug Interacti project.
- 7. Madeline Aranda, the ACME Secretary, transferred to the Financial Aids Office.

The balance of the staff will likely be assigned either to new computing facilities within the Medical Center or to other service computing facilities on the Stanford campus.

CURRENT ACME PERSONNEL

V. PROJECT DESCRIPTIONS

CORE RESEARCH & DEVELOPMENT

A. DENDRAL Project: DENDRAL Project: DENDRAL Realtime

Investigator: Edward Feigenbaum, Joshua Lederberg, and Carl Djerassi

Dept. of Chemistry, Computer Science, and Genetics

The DENDRAL project involves collaboration between the Instrumentation Research Laboratory operating under NASA grant NGR-05-020-004, investigators operating under NIH grant $RROO612$, and $ACME$.

The emphasis of the DENDRAL-ACME efforts is computer science, while that of IRL-ACME endeavors is data acquisition and computer instrument control.

The DENDRAL project aims at emulating in a computer program the inductive behavior of the scientist in an important but sharply limited area of science; organic chemistry. Most of the work is addressed to the following problem; given analytic data (the mass spectrum) of an unknown compound, infer a workable number of plausible solutions, that is, a small list of candidate molecular structures. In order to complete the task, the DENDRAL program then deduces the mass spectrum predicted by the theory of mass spectrometry for each of the candidates and selects the most productive hypothesis, i.e., the structure whose predicted spectrum must closely matches the data.

The project has designed, engineered, and demonstrated a computer program that manifests many aspects of human problem solving techniques. It also works faster than human intelligence in activing problems chosen from an appropriately limited domain of types of compounds, as illustrated in the cited publications.

Some of the essential features of the DEXDRAL program include:

Conceptualizing organic chemistry in terms of topological graph theory, i.e., a general theory of ways of combining atoms.

Embodying this approach in an exhaustive HYPOTHESIS GENERATOR. This is a program which is capable, in principle, of "imagining" every conceivable molecular structure.

Organizing the GENERATOR so that it avoids duplication and irrelevancy, and moves from structure to structure in an orderly and predictable way.

Core Research & Development (Continued)

The key concept is that induction becomes a process of efficient selection from the domain of all possible structures. Heuristic search and evaluation are used to implement this "efficient selection."

Most of the ingenuity in the program is devoted to heuristic modifications of the GENERATOR. Some of these modifications result in early pruning of unproductive or implausible branches of the search tree. Other modifications require that the program consult the data for cues (pattern analysis) that can be used by the GENERATOR as a plan for a more effective order of priorities during hypothesis generation. The program incorporates a memory of solved sub-problems that can be consulted to look up a result rather than compute it over and over again. The program is aimed at facilitating the entry of new ideas by the chemist when discrepancies are perceived between the actual functioning of the program and his expectation of it.

The DENDRAL research effort has continued to develop along several dimensions during Fiscal 1973. The mass spectra of some previously uninvestigated compounds were recorded, The computer program has been extended to analyze the mass spectra of a more complex class of compounds, using new kinds of data. The artificial intelligence work on theory formation and program generality has also progressed.

The techniques of artificial intelligence have been applied successfully for the first time to a problem of direct biological relevance, namely the analysis of the high resolution mass spectra of estrogenic steriods. The performance of this program has been shown to compare favorably with the performance of trained mass spectroscopists. (see Smith, et al. (1972)

Of particular significance in this effort were, in addition to exceptional performance, the potential for analysis of estrogens without prior separation, and for generalization of the programming approach to other classes of molecules.

Because of the structure of the Heuristic DENDRAL program for estrogens, it is immaterial whether the spectrum to be analyzed is derived from a single compound or a mixture of compounds. Each component is analyzed, in terms of molecular structure, in turn, independently of the other components. This facility, if successful in practice, would represent a significant advance of the technique of mass spectrometry. Many problem areas, because of physical characteristics of samples or limited sample quantities, could be successfully approached utilizing the spectra of the unseparated mixtures. Even in combined gas chromatography/mass spectrometry (GC/MS), many mixture components will be unresolved and an analysis program must be capable of dealing with these mixtures.

We have, in collaboration with Prof. H. Adlercreutz of the University of Helsinki, recently completed a series of analyses of various fractions of estrogens extracted from body fluids and supplied to us by Prof. Adlercreutz. These fractions (analyzed by us as unknowns) were found to contain between one and four major components, and structural analysis of each major component was carried out successfully by the above program. These mixtures were analyzed as unseparated, underivatized compounds. The implications of this success are considerable. Many compounds isolated from body fluids are present in very small amounts and complete separation of the compounds of interest from the many hundreds of other compounds is difficult, time-consuming and prone to result in sample loss and contamination. We have found in this study that mixtures of limited complexity, which are difficult to analyze by conventional GC/MS techniques without derivatization (which frequently makes structural analysis more difficult), can be rationalized even in the presence of significant amounts of impurities. A manuscript on this study has been submitted to the Journal of the American Chemical Society.

In the past year we have extended our library of high resolution mass spectra of estrogens to include 67 compounds. These data represent an important resource and have been included (as low resolution spectra for the moment) in a collection of mass spectra of biologically important molecules being organized by Prof. S. Markey at the University of Colorado.

The Heuristic DENDRAL program for complex molecules has received considerable attention during the last year in order to remove compound class specific information or program strategies. By removing information which is specific to estrogens, the program has become much more general. This effort has resulted in a production version of the program which is designed to allow the chemist to apply the program to the analysis of the high resolution mass spectrum of any molecule with a minimum of effort. Given the spectrum of a known or unknown compound, the chemist can supply the following kinds of information to guide analysis of the mass spectrum: a) Specifications of basic structure (superatom) common to the class of molecules. b) Specification of the Fragmentation rules to be applied to the superatom, in the form of bond cleavages, hydrogen transfers and charge placement. c) Special rules on the relative importance of the various fragments resulting from the above fragmentations. d) Threshold settings to prevent consideration of low intensity ions. e) Available metastable ion data and the way these data are subsequently used -- to establish definitive relationships between fragment ions and their respective molecular ions. f) Available low ionizing voltage data - to aid the search for molecular ions. κ) Results of deuterium exchange of labile hydrogens $-$ to specify the number of, e.g., $-$ OH groups.

Core Research & Development (Continued)

We have been very successful in testing the generality of the program, with particular emphasis on other classes of biologically important molecules. We have used the program in analysis of high resolution mass spectra of progesterone and some methylated analogs, a small number of androstane/ testosterone related compounds, steroidal sapogenins and n-butyl-trifluoroacetyl derivatives of amino acids.

The Heuristic DENDRAL performance program described above is an automated hypothesis formation program which models "routine", day-to-day work in science. In particular, it models the inferential procedures of scientists identifying components, such as those found in human body fluids. The power of this program clearly lies in its knowledge about various classes of compounds normally found in body fluids, which knowledge allows identification of the compounds.

The Meta-DENDRAL program described in this part is a critical adjunct to the performance program because it is designed to supply the knowledge which the performance program uses. Theory formation is essential in order to carry out the routine analyses - either by hand or by computer. However, the staggering amount of effort required to build a working theory (even for a single class of compounds) holds back the routine analyses. The goal of the Meta-DENDRAL program is to form working theories automatically (from collections of experimental data) and thus reduce the human effort required at this stage. By speeding up the time between collecting data for a class of compounds and understanding the rules underlying the data, the Meta-DENDRAL program will thus provide an improvement in the development of diagnostic procedures.

Detailed accounts of this research are available in the DENDRAL Project annual report to the National Institutes of Health, in several papers already published, and in manuscripts submitted for publication.

- 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, II. 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, 0. 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).

Core Research & Development (Continued)

- 6. H. S. Hertz, R. A. Hites and K. Biemann, Analytical Chemistry, 43, 681 (1971), S. L. Grotch, ibid., 43, 1362 (1971).
- 7. E. A. Feigenbaum, B. G. Buchanan, and J. Lederberg, 'On Generality and Problem Solving: A Case Study Using the DENDRAL Program". In Machine Intelligence 6 (B. Meltzer and D. Michie, eds.) Edinburgh University Press (1971). (Also Stanford Artificial Intelligence Project Memo No. 131.)
- a. A. Buchs, A. B. Delfino, C. Djerassi, A. M. Duffield, B. G. Buchanan, E. A. Feigenbaum, J. Lederberg, G. Schroll, and G. L. Sutherland, "The Application of Artificial Intelligence in the Interpretation of Low-Resolution Mass Spectra', Advances in Mass Spectrometry, 5, 314.
- 9. B. G. Buchanan and J. Lederberg, "The Heuristic DENDRAL Program for Explaining Empirical Data".. In proceedings of the IFIP Congress 71, Ljubljana, Yugoslavia (1971). (Also Stanford Artificial Intelligence Project Memo No. 141.)
- 10. B. G. Buchanan, E. A. Feigenbaum, and J. Lederberg, "A Heuristic Programming Study of Theory Formation in Science.' In proceedings of the Second International Joint Conference on Artificial Intelligence, Imperial College, London (September, 1971). (Also Stanford Artificial Intelligence Project Memo No. 145.)
- 11. Buchanan, B. G.,Duffield, A. M.,Robertson, A. V., "An Application of Artificial Intelligence to the Interpretation of Mass Spectra", Mass Spectrometry Techniques and Appliances, Edited by George W. A. Milne, John Wiley & Sons, Inc., 1971, p. 121-77.
- 12. D. H. Smith, B. G. Buchanan, R. S. Engelmore, A. M. Duffield, A. Yeo, E. A. Feigenbaum, 3. Lederberg, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference VIII. An approach to the Computer Interpretation of the High Resolution Mass Spectra of Complex Molecules. Structure Elucidation of Estrogenic Steroids", Journal of the American Chemical Society, 94, 5962-5971 (1972).
- 13. B. G. Buchanan, E. A. Feigenbaum, and N. S. Sridharan, "Heurist Theory Formation: Data Interpretation and Rule Formation". In Machine Intelligence 7, Edinburgh University Press (1972).
- 14. Brown, H., Masinter L., Hjelmeland, L., "Constructive Graph Labelin Using Double Cosets". Discrete Mathematics (in press), (Also Computer Science Memo 318, 1972.)
- 15. B. G. Buchanan, Review of Hubert Dreyfus' "What Computers Can't 1973). (Also Stanford Artificial Intelligence Project Memo No. 181) : A Critique of Artificial Reason", <u>Computing Reviews</u> (January
- 16. D. H. Smith, B. G. Buchanan, R. S. Engelmore, H. Aldercreutz and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference IX. Analysis of Mixtures Without Prior Separation as Illustrated for Estrogens". Submitted to the Journal of the American Chemical Society.
- 17. D. H. Smith, B. G. Buchanan, W. C. White, E. A. Feigenbaum, C. Djerassi and J. Lederberg, "Applications of Artificial Intelligence for Chemical Inference X. Intsum. A Data Interpretation Program as Applied to the Collected Mass Spectra of Estrogenic Steroids". To be submitted.

The preceding comments on DENDRAL involve Parts A and C as described in the table below. The balance of this section deals with Part B, instrumentation aspects.

Part A: Applications of Artificial Intelligence to Mass Spectrometry. Part B(i): Mass Spectrometer Data System Development. Part B(ii): Analysis of the Chemical Constituents of Body Fluids. Part C: Extending the Theory of Mass Spectrometry by Computer.

ACME computer support for DENDRAL Part B has been treated as ACME core research activity during FY73. Excerpts from DENDRAL's annual report follow, detailing recent accomplishments.

The large volume of data which must be reduced and interpreted from each GC/MS analysis of a body fluid sample together with the increasing number of samples which must be processed to be responsive to clinical needs , point to more and more highly automated and reliable GC/MS systems. This portion of the proposal addresses the problems of developing and applying such automated systems from several points of view. First, we propose to investigate the integration of sophisticated computer analysis programs into data reduction, data interpretation, and instrument management functions in order to progressively relieve the chemist from manually performing these tasks. Second, we will maintain the daily operation of our GC/MS systems for the on-going investigation of clinical applications and the acquisition of data necessary for the development of automated interpretation programs.

Cur overall objectives for automating GC/MS systems comprise a number of specific subgoals including a) implementing highly automated and reliable systems for the acquisition and reduction of low resolution, high resolution, and metastable mass spectral data; b) implementing a data system to support combined gas chromatography/high resolution mass spectrometry; c) automating the location and identification of constituents of body fluid extracts from gas chromatogram and mass spectrum information for the routine application of these techniques to clinical problems; and d) investigating the intelligent closed loop control of mass spectrometer systems in order to optimize the data acquired relative to the task of data interpretation.

Core Research & Development (Continued)

A. Mass Spectrometer Data System Automation

Concentrating initially on the MAT-711 spectrometer, we have made significant progress toward a reliable, automated data acquisition and reduction system for scanned low and high resolution spectra. This system is largely failsafe and requires no operator support or intervention in the calculation procedures. Output and warnings to the operator are provided on a CRT adjacent to the maas spectrometer. The system contains many interactive features which permit the operator to examine selected features of the data at his leisure. The feedback currently provided to the operator to assist in instrument set-up and operation can just as well be routed to hardware control elements for these functions thereby allowing computer maintenance of optimum instrument performance.

Progress in this area is an integration of our efforts in hardware and software improvements:

HARDWARE - The basic system consists of the mass spectrometer interfaced to a PDP-11/20 computer for data acquisition, pre-filtering, and time buffering into the ACME time-shared 360/50. The more complex aspects of data reduction are done in the 360/50 since the PDP-11 has limited memory and arithmetic capabilities. New interfaces for mass spectrometer operation and control have been developed. The interfaces can handle (through an analog multiplexer) several analog inputs and outputs which require that the PDP-11 computer be relatively near the mass spectrometer. We now have the capability for the following kinds of operation through the new interfaces.

- i) Computer selection of digitization rate.
- ii) Computer selection of data path (interrupt mode or direct memory access (DMA).
- iii) Direct memory access for faster operation in the data acquisition mode.
- iv) Computer selection of analog input and output channels.
- $v)$ Sensing of several analog channels through a multiplexer (e.g., ion signal, total ion current).
- vi) Magnet scan control. This control can be exercised manually or set by the computer. It controls both time of scan and flyback time. Coupled with selection of scan rate, any desired mass range can be scanned at any desired scan rate.
- vii) The computer can monitor the mass spectrometer's mass marker output as additional information which will be used to effect calibration.
$SOFTWARE - Automatic instrument calibration and data reduction pro$ grams have been developed to a high degree of sophistication. We can now accurately model the behavior of the MAT-711 mass spectrometer over a variety of scan rates and resolving powers. Our instrument diagnostic routines are depended upon by the spectrometer operator to indicate successful operation or to help point to instrument malfunctions or setup errors. Some features of these programs are described below.

- i) Data Acquisition. Programs have been written which permit acquisition of peak profile data at high data rates using the PDP-11 as an intermediate data filter and buffer store between the mass spectrometer and ACME. This allows data acquisition to proceed even under the time constraints of the time-sharing system. Storage of peak profiles rather than all data collected has greatly reduced the storage requirements of the program and saves time as the background data (below threshold) are removed in realtime. An automatic thresholding program is in operation which statistically evaluates background noise and thresholds subsequent data accordingly. Amplifier drift can thus be compensated. We have developed some theoretical models of the data acquisition process which suggest that high data acquisition rates are not necessary to maintain the integrity of the data. Demonstration of this fact with actual data has helped relieve the 'burden of high data rates on the computer system, particularly as imposed by GC/MS operation, and permits more data reduction to be accomplished in realtime or alternatively reduces the required data acquisition computer capacity.
- ii) Instrument Evaluation. A high resolution mass spectrometer operating in a dynamic scanning mode is a complex instrument and many things can go wrong which are difficult for the operator to detect in realtime. In order for the computer to assist in maintaining data quality, it must have a model of spectrometer operation on the basis of which data quality can be assessed and processing suitably adapted as well as instrument performance optimized. We have developed a program which monitors the state of the mass spectrometer.
- iii) Data Reduction. A program has been written which allows automatic reduction of high resolution data based on the results of the prior instrument evaluation data. Conversion of peak positions in time to the corresponding mass values is effected by mapping each spectrum into the calibration model developed previously. The interpolation algorithm between reference calibration points incorporates

a quadratically varying exponential time constant to account for the second order character of a magnet discharging through a resistance and a capacitance as well as an offset at infinite time to account for residual magnetization affecting accuracy at low masses.

Perfluorokerosene (PFK) peaks, introduced into high resolution mass spectra for internal mass calibration, are distinguished from unknown peaks by a pattern recognition algorithm which compares the relationships between sequences of reference peaks in the calibration run with the set of possible corresponding sequences in the sample run. The candidate sequence is selected which best approximates calibrated performance within constraints of internally consistent scan model variations. This approach minimizes the need for selection criteria such as greatest negative mass defect for reference peaks, the validity of which cannot be guaranteed. Excellent performance results from using sequences containing 10 reference peaks.

Unresolved peaks are separated by a new analytical algorithm, the operation of which is based on a calculated model peak derived from known singlet peaks rather than the assumption of a particular parametric shape (e.g., triangular, Gaussian, etc.) This alogorithm provides an effective increase in system resolution by a factor of three thereby effectively increasing system sensitivity. By measuring and comparing successive moments of the sample and model peaks, a series of hypotheses are tested to establish the multiplicity of the peak, minimizing computing requirements for the usually encountered simple peaks. Analytic expressions for the amplitudes and positions of component peaks have been derived in the doublet case in terms of the first four moments of the peak complex. This eliminates time consuming iteration procedures for this important multiplet case. Iteration is still required for more complex multiplets.

Elemental compositions are calculated from high resolution mass values with a new, efficient table look-up algorithm developed by Lederberg.

Future work will extend these ideas to a system for the acquisition of selected metastable information as well as to include the quadrupole system used in the routine low resolution clinical work.

B. GAS Chromatography/High Resolution Mass Spectrometry.

We have recently verified the feasibility of combined gas chromatography/high resolution mass spectrometry (GC/HRMS). Using the programs described above we can acquire selected scans and reduce them automatically,

although the procedures are slow compared to "realtime" due to the limitations of the time-shared ACME facility. We have recorded sufficient spectra of standard compounds to show that the system is performing well.

We have begun to exercise the GC/HRMS system on urine fractions containing significant components whose structures have not been elucidated on the basis of low resolution spectra alone. Whereas more work is required to establish system performance capabilities, two things have become clear: 1) GC/HRMS will be a useful analytical adjunct to our low resolution GC/MS clinical studies to assist in the identification of significant components whose structures are not elucidated on the basis of low resolution spectra alone, and 2) the sensitivity of the present system limits analysis to relatively intense GC peaks.

Recent experiments in operation of the mass spectrometer in conjunction with the gas chromatograph have also shown that the present ACME computer facility cannot provide the rapid service required to acquire repetitive scans at either high or low resolving powers. We can, however, acquire scans on a periodic basis, meaning most GC peaks in a run can be scanned once at high resolving power. We are presently implementing a disk on the PDP-11 to act as a temporary data buffer between the mass spectrometer and ACME. This disk will allow acquisition of repetitive scans, while data reduction must be deferred to completion of the GC run.

C. Automated GC/MS Data Reduction

The application of GC/MS techniques to clinical problems as described in Part $B(i)$ of this proposal has made clear the need for automating the analysis of the results of a CC/MS experiment. Previous paragraphs dealt with the problems of reducing raw data in preparation for analysis. At this point the data must be analyzed with a minimum of human interaction in terms of locating and identifying specific constituents of the GC effluent. The problem of identification is addressed by the library search and DENDRAL mass spectrum interpretation programs discussed in Part A of this proposal. The problem of locating effluent components in the GC/MS output involves extracting from the approximately 700 spectra collected during a GC run, the 50 or so representing components of the body fluid sample. The raw spectra are in part contaminated with background "column bleed" and in part composited with adjacent constituent spectra unresolved by the GC.

We have begun to develop a solution to this problem with very promising results.

Core Research & Development

D. Closed-Loop Instrument Control.

The task of collection of different types of mass spectral information (e.g., high resolution spectra, low ionizing voltage spectra snd selected metastable information) under closed loop control during a GC/MS experiment is extremely difficult and may not be realizable with current technology. We are studying this problem in a manner which will allow the system to be used for important research problems (e.g., routine analysis of urine fractions without fully closed loop control) while aspects of instrument control strategy are developed in an incremental fashion.

B. Time Oriented Database System (TCP)

Investigator: Dr. James Fries and the ACME Staff Pro.iect: J FRIES . DATABANK F GERMAN. TOD DATABANK.TODD

Dept. of Medicine - Immunology and ACME

In 1970 and 19'71, Dr. James Fries in the Division of Immunology of the Department of Medicine developed concepts and implemented programs which he labeled "Time Oriented Database". One of the first steps was the development of standard forms for use in the medical record. These forms are completed manually and require no computer intervention or interaction. Use of the new medical record forms has proved highly desirable in several clinics at Stanford since that time, with or without the associated computer programs. The relationship of the computer to the project makes possible rapid comparison and statistical analysis of various data items covering multiple visits for one patient or for many patients.

In the summer of 1972, a design study was completed which would generalize the use of the TOD programs on the ACME system so that several divisions could use a common set of programs. The design effort was handled primarily by Stephen Weyl with assistance from Gio Wiederhold and Frank Germano. Implementation of the new generalized TOD programs was managed by Frank Germano with Stephen Weyl, Rick Giusti, Bob Bassett, and Jane Whitner handling the programming.

As of May 1, 1973, several TOD databanks had been implemented and several more had been planned. The table below reflects the progress to that date.

TOD Implementation Progress Report (May 1, 1973)

PRESENT TOD DATABANKS

TOD DATABANKS GOING THROUGH DEFINITION PROCESS

GROUPS CONSIDERING A TOD DATABANK DEFINITION

- Dr. Wilbur Childrens Hospital
- Dr. Miller Childrens Hospital
- Dr. V. Johnson Pediatrics
- Dr. A. Hackel
- Dr. M. Bagshaw Radiology

The system which was announced in January 1973 is but a first step in development of database systems at Stanford, Clearly more development effort will follow which will improve the data entry techniques to be employed, enhance quality control of data entered, and increase the amount of shared data in the files.

The following pages contain four ACME Notes written to document the TOD system, along with explanatory remarks. The four ACME Notes are:

> TODI \cdot - Introduction to TOD System TODREF - Index to TOD ACME Notes TODDDL - TOD Databank Description Language TODCST - Analyzing the Costs of Running a TOD Databank

The TOD system is a set of programs available from ACME designed to aid users in the creation, maintenance, and use of computer "databanks" which store patient-related information over time. These programs are available as TOD publi programs on the PL/ACME system. If a user can conceptually view his patier data in the form of a three-dimensional array, indexed by patient, parameter, and time, he can use the TOD system. A recently conducted database review of medical data stored on the ACME computer system and other Stanford computers revealed that many of these databanks have this form. The results of this survey are summarized in ACME Note DBS.

Flexibility and Independence

In order to offer a system of programs which support most patient-related databases implemented on the ACME facility, a large degree of flexibility and independence had to be built into the system. The TOD approach is a decentralized one, in which each division maintains a separate databank, whose inter-relation to all databanks is well defined. Each TOD databank is set up and used under one ACME name and project. The databank planner is the administrator of that databank, not ACME. To provide for user definition, an extra file is added to the databank. This file is called a SCHEMA file; it describes the form of the databank. It stores that information which makes each TOD databank unique for its user. Public programs which act on a TOD databank look to this file for descriptive information about the databank. This information is then used by the various programs which act on the databank during their operation.

Advantages of TOD for A User

Use of the TOD system can offer several advantages to the user. Some of the direct advantages are discussed below.

1. Less Effort to Utilize a Patient-Related Databank:

Prior to TOD each user essentially had to write programs to set up, maintain, and use a databank. This represented a great duplication of effort. The databanks tended to be implemented according to different, rather arbitrary conventions. Moreover, because of the diversity of form for the various databanks, sharing of information could only be done on a case-bycase basis and with special programming. Use of TOD will reduce programming effort for users who store patient-related data.

2. Data Sharing

Because of the existence of the SCHEMA file, all the information required to allow sharing of data is in one place and in computer-readable form. This will allow data sharing between TOD users to occur more easily in the future.

3. High-Level Documentation

Aside from providing information to programs which operate on the TOD system, the SCHEMA file provides information to programmers and users of a databank. This information, describing individual items in the databank, is defined by a Schema Language called DDL (Database Description Language) which uses a PL/ACNE-like syntax for its declarations. This common language forms the basis for unambiguous communication among TOD databank groups. This communication process is strengthened by the fact that the different groups share a common general core set of programs and a common general file structure. Details of an individual databank are described using the Schema Language.

4. Operational Statistics

All the TOD programs store statistics which describe the operation of the databank. Careful review of these statistics in conjunction with the monthly summary of ACME charges will give the user a much clearer picture of what his computer dollar is buying.

5. Common Improvements

As ACME and users find ways to improve the TOD system programs and procedures in terms of capabilities and cost-effectiveness, these improvements will be passed along to TOD users by changes in the TOD programs and systems documentation to be implemented by means of monthly "releases" of the system. These releases will be upwards compatible. If a user writes a specialized program which he feels is worthy of sharing with the TOD group, this program can easily be generalized and made available as part of the TOD system.

Overview of the TOD System

The programs comprising the TOD system fall into four groups: data entry, data update, data retrieval and analysis, and TOD System Utility Programs. Figure I summarizes these groups.

DATA UPDATE

SYSTEM UTILITY: TO CSTOK, TO TRANS, TD INDEX, TD RANGE, TD TRA. TD CHECK

TOD Programs

The programs named in Figure I are representative of the programs that will comprise the initial TOD system. Most of these are main programs, although a few are subprograms which can be included in a user-written program. Table I below summarizes the purpose of these programs.

Table I. TOD System Programs

ACME statistical programs. (Not implemented.)

TOD Files

A TOD databank contains a number of inter-related files. Four of these files are required: td schem, td desc, td head, and td parm. In addition to these files, several auxiliary files can be added to the system to make retrieval of certain information faster. These files are td index, td range, td htpse, and td_ptpse. Table II summarizes the purposes of these file

Table II. TOD Files

File Information Content

- td schem Description of the databank in PL/ACME DECLARE-type statements.
- td-desc Internal form of the databank description.
- td head Demographic patient information. One record per patient.
- td parm Information measured at a point in time. One record per time per patient.
- td index (HEADER item value, KEY to HEADER file) pairs sorted on header values. One such group for each header element that is indexed.
- td_range For each patient, the hi and lo ranged parameter values across all parameter records associated with the patient over time. Only those parameters which are ranged are included.

td htpse td-ptpse For TRAXSPCSED data items these files contain the same information as the HEADER and PARAMETER files except that the ordering is such that oll valmes for a particular item are <u>contiguous</u>, making questicus
which relate to specific foems much faster to answer.

The Purpose of the Present TOD Effort

The present TOD implementation is not to be the system to end all information retrieval systems. Its capabilities have been limited in order to assure that a demcnstratable working system can be swiftly implemented. Nevertheless, a full set of capabilities are provided to handle most of the users who are following patients over time. Once a number of TOD users exist, who speak a common language, further extensions to the system can be planned in a meaningful manner.

ACME views the TOD system as a set of programs which allows users who follow patients over time to set up, maintain, and use a databank in a simple and efficient manner. The present TOD effort is a study of the patient databank question in the Stanford Medical Center.

Further Reading

A reference to all ACME notes describing the operation and use of various portions of the TOD system and its implementation is given in ACME Note TODREF. ACME Note

Index to TCD ACME Notes

TODREF-1 Steve Weyl April 4, 1973

This note is a comprehensive index to the set of ACNE Notes describing the TOD (Time-Oriented Databank) system, The index is given in three parts: Part I references the notes that all planners and users should be familiar with. Part II references the file structure and system implementation notes, which are primarily of interest to systems analysts and programmers. Part 111 references historical notes, notes describing administrative procedures for the TOD system, and notes associated with individual TOD databanks.

The TOD system is a set of programs available from ACME designed to aid users in the creation, maintenance, and use of computer "databanks" which store patient-related information over time.

ACME noteL TODI and TDOV give an overview of the TOD system. ACME note TODD is the original design document for TOD and is primarily of historical interest, since many of the conventions suggested there have been modified in the course of implementation.

* Notes marked with an asterisk $(*)$ had not yet been published at the time this issue of TODREF went to press.

PART $I - USE OF THE TOD SYSTEM$

A. General Introduction and Overview

TDOV TOD System Overview -- F. Germano

B. Planning and Defining a Databank

- TDPRE Redefinition of a TOD Databank Using TD-RECOM -- S. Weyl
- C. Entering and Correcting Data
	- *TDUB How to Enter Data on TOD -- V. Wiederhold
	- TODPDG Checking Data Values and File Linkage Using Program TD-CHECK -- R. Giusti
- D. Report Generating Programs

E. Retrieval and Analysis Programs

- TODPDD TOD Retrieval Module Summary Sheet -- F. Germano
- TODPDO Definition of Patient Subsets for Analysis Using Programs TD SALL, TD SAND, TD SOR, and TD SSUPR -- S. Weyl
- TODPDB TOD Scatterplot Program -- F. Germano
- TODPDC TOD Reviewdx Program -- F. German0
- TODPDE TOD Survival Kit - User Instructions --J. Whitner
- TODPDJ TOD Debug Lister Program TD-QKLST -- R. Giusti
- TODPDM Using TOD Retrieval Nodules as Debug Programs -- R. Giusti
- *TODSUR TOD Survival Kit Computational Methods -- M. Hu

F. TOD Utility Programs

TODPDG Checking Data Values and File Linkage Using Program TD_CLECK -- S. Weyl, R. Giusti

- S. Weyl, R. Giusti
- TODPDK Constructing TOD Index Files with Program TD-INDEX -- R. Giusti
- G. Writing Your Own Analysis Programs

TIDA TOD Analysis Programs -- F. Germano

H. Operational Costs of TOD Databanks

- TODPDA Operational Overview for a TOD Databank -- F. German0
- TODCST Analyzing the Costs of Running a TOD Databank -- F. German0

PART II - INTERNAL DOCUMENTATION

A. Program Documentation

B. File Structure

- TIDG Record l in the TOD Descriptor File, td-desc --F. German0
- TIDH Structure of the Subset Library File, td subs -- S. Weyl

PART III - OTHER ACME NOTES

A. Historical

- TODD Definition of the PL/ACME Time-Oriented Databank Protocol -- S. Weyl
- DBT ACME Data Base for Cancer Virus Tumor Samples (Medical Microbiology - Dr. Hayflick) -- S. Weyl
- DBD ACME Data Bases for Drs. Eugene Dong and Phillip Caves - Cardiovasular Surgery Research -- S. Weyl
- MOP Comment on Medical Applications Oriented Preliminary Data Base -- S. Weyl
- PMOD Need for a Medical Applications Oriented Data Base Protocol and Support Facility -- S. Weyl
- BSPD Sharing Patient Data Files -- G. Wiederhold
- DBS Present and Potential Patient-Related Databanks at the Stanford Medical Center -- F. Germano, G. Wiederhold
- HTP Preliminary Data Base for heart Transplant Pilot Research on Dogs $--$ S. Weyl

B. TOD Administrative Procedures

- TODADM Administrative Procedures for the PL/ACME Time-Oriented Databank (TOD) -- F. Germano, S. Weyl
- C. Notes on Individual TOD Databanks
	- TDUONA Programs PRELET ONCOLET: Oncology Letter Writing Programs -- J. Whitner
	- *TDUONB Time-Oriented Databank for the Oncology Clinic -- S. Weyl

D. Keyword Index to TOD Notes

*TODIDX Keyword Index to TOD Notes -- F. Germano, S. Weyl

OTHER REFERENCES

- 1. Wiederhold, Gio, An Advanced Computer System for Medical Research, PROCEEDINGS OF THE IBM JAPAX COMPUTER SCIENCE SYMPOSIUM--Research and Development and Computer Systems
- 2. Frey, Girardi, Wiederhold, A Filing System for Medical Research, BIOMEDICAL COMPUTING, (2) (1971).
- 3. Wiederhold, Gio, Database Structures and Schemas (to be published)
- 4. Fries, James, Time Oriented Medical Research and a Computer Data Bank, JAMA, vol. 222, no. 12, Dec. 18, 1972, pp. 1536-1542.

Dist: Staff/TOD/All

The TOD Databank Description Language

The TOD Databank Description Language is a means to define medical data in a less ambiguous form than has been used in the past. ACME Note TODDDL describes this defining capability.

In order to make the task of defining a patient databank easier, forms were designed which contained spaces for the same information required by the TOD Databank Description Language. A sample TOD Databank Element Definition form appears on the next page.

Once several databanks were defined using the TOD Databank Description Language, the concept of the TOD Data Descriptor Dictionary came into being. The Stanford Medical Center Data Descriptor Dictionary is a listing of the data elements in all the TOD databanks. This listing is arranged in order by the symbolic (short 8 -character) name assigned to TOD data elements by individual databank planners. To each symbolic name a two-character suffix has been appended to indicate which TOD databank the data element resides in. When several databanks have the same symbolic name for a data item (which should only occur when the elements are indeed the same data variable in each of the individual databanks), they appear together in the listing, each with its own unique suffix.

The data dictionary pulls together in one place the variables stored in all the TOD databanks. It enables new databank planners to see what data already exists in other TOD databanks, but more importantly, it shows what conventions, such as data checking or units, were assigned to the data items.

A sample page for the TOD Data Descriptor Dictionary follows.

STANFORD TOD DATABANK ELEMENT DEFINITION

TOD DATABANK

SAMPLE
ELEMENT
FORM

DEFINITION

 \mathbf{F} ЪÓ \blacksquare

SAMPLE PAGE FROM
TOD DATA DESCRIPTOR DICTIONARY

 $\pmb{\cdot}$ $\tilde{5}$ \mathbf{I}

 \bar{z}

 $\Gamma\text{-}\mathbf{A}\mathbf{T}\mathbf{A}\text{-}\mathbf{I}$

The TOD Databank Description Language

A. The TOD Schema and DDL:

A TOD databank is constructed and accessed according to a carefully defined databank description called a schema. The schema is specified by a set of 'high level" language statements which are stored on a textfile, much like a PL/ACKE program. These statements are translated into an internal form by the schema translation program, TD_TRA (ACME Note TDPT).

The high level schema language will henceforth be referred to as the databank description language, DDL. The syntax of DDL resembles PL/ACME syntax, except that it only provides for declarations. DDL is designed to require accuracy and completeness of definition, since these are essential for effective databank operation.

The schema serves two important functions: First, its DDL representation gives an explicit documentation of the content, unit of measurement, reference name, and type of each data item; and it indicates the data initialization, range checking, encoding, and retrieval file maintenance which must be performed. Second, the internal form of the schema provides generalized TOD programs with information about the structure of a specific user's databank (see example in ACME Note TODD, Part III, Section A.2.d.i.).

The necessary information for writing a DDL schema can be accumulated and proofed on a convenient printed form, copies of which may be obtained at the ACME office.

B. Basic Semantics of DDL:

A "single piece" of information is stored in a data item or just item. The information stored in an item is referred to as the value of that item.

The items of a databank can be partitioned into two general categories. Items in the first category are recorded only once and they store demographic information or background information associated with a patient. Items in the second category store the numeric values of several time-dependent parameters recorded for each patient-visit. A 'visit" corresponds to a physician interview, or any other point in time at which information about a patient is logged.

DDL represents the two categories of data items as formal arrays. Each formal array element corresponds to a single data item, so that an array of elements is a list of related items. The category of demographic items is represented by the HEADER formal array. The category of time-dependent items is represented by the PARAMETER formal array.

The size of each formal array, and thus the size of each category of items, is established with a declaration statement. Then each formal array element may be assigned a 5-tuple of attributes. These attributes describe the data item associated with a formal array element. The attributes define content, a measurement unit, name, data type. and retrieval type for an item. Unassigned formal array elements are place-holders expediting the introduction of new data items into the databank.

The INITIAL statement allows a user to define the initial (i.e. default) value of each item prior to data entry. Careful assignment of default values can result in major savings of secretarialand processing time. Also, by reducing the amount of data which must be entered per patient-visit, the data entry error rate will be lowered.

```
Formal Syntax for DDL:
```
DDL sentences

comments */ $/$ *

HEADER DECLAF... PARAMETER (array size) : H \overline{P}


```
INITIA . name [VALUE(initial value)][SAME];
```
For explanation of formal syntactic conventions, see ACME Note WAA.

Explan tion of Terms:

Capita ized words are keywords recognized independently of capitalization. Lower case words are explained below.

"comments" is a string of any characters. "array size", "array elt no"(array element number), "string length", "min", & "max"are numbers. "content" and "unit" are literal strings enclosed in single quotes. "name" is a PL/ACME variable name. "initial value" may be literal string (enclosed in single quotes) or a number.

Textfile Format:

Multiple sentences may occur on a single textfile line. Extra non-imbedded blanks are ignored. (Note: Comments are not permitted inside of sentences.)

Note o: Semantics:

Some symbactically correct sentences are not semantically acceptable, and will be rejected by the translator. See ACHENote TDPT for these exceptions.

 \mathbf{r} $\vec{5}$

D. Detailed Semantics for DDL Statements :

Comments, enclosed in the PL-conventional " $/*$ " and "* $/$ " are ignored by the schema translator and serve only as documentation.

A declaration statement specifies the number of elements in each formal array. Before data items can be assigned attributes, their associated formal arrays must be declared of adequate size. $"H"$ and "P" are recognized abbreviations for "HEADER" and "PARAMETER".

The formal assignment statement defines for formal array elements the attributes of their associated data items. There are five attributes which must all be specified. We will consider each separately.

The first attribute of a data item is a character string of length \leq 40 describing the content of that data item. For example, if a HEADER item stores the address of the referring physician for each patient,its content might be described as 'ADDRESS OF REFERRIIiG PHYSICIAN'. The content attribute is included in the schema for documentation purposes. It may eventually be used for computer composition of Time Oriented Record forms.

The second attribute of a data item is a character string of lengt \leq 10 describing a standard unit of measurement. For example, if a PARAMETER item stores the white blood cell count for each patientvisit, its units might be described as 'x1000'. Aunit must always be specified and a null response will not be accepted. For the HEADER item suggested in the first example, itsunit would have to be specified as 'address' or 'none'. The reason for always requiring units is to assure that databank planners fully specify the meaning of each data value, so that data can be shared.

The third attribute of a data item is a (short) name by which the data item may be uniquely identified. This name must be a valid PL/ACME variable name. Names will always be recognized irrespective of capitalization. In the second example above, "WBC" might name "white blood cell count".

Names are an important symbolic attribute of data items, and the use of standard names will greatly simplify communication of information between two data banks and merging of data banks. To assist in the standardization of data item names, a sorted list of names and the attributes of their associated data items will be maintained by the TOD manager.

Public databank procedures will make use of certain standard names for automatic update of data values. For example, the data item named "age" will automatically have the current age of a patient stored in it (as computed from the "date" for this visit and the "birthdat" item for this patient). These standard items are defined in Section E.

In the INITIAL statement, HEADER or PARAMETER elements are identified by their names.

Names will be used in tabular output programs for column headings.

The fourth attribute of a data item specifies the data type of information stored in this item. The following types are available:

- 1) VALUE data items can meaningfully assume continuous numeric values with six significant figures. They are stored as ACME single precision floating point numbers. VALUE data items are assumed by analysis programs to be measured on an interval scale.
- 2) +RANGE data items can assume the values 0, 1, 2, 3, 4 (of the "+RANGE scale"), which is treated by analysis programs as an ordinal scale.
- 3) DISCRETE data items can only meaningfully assume discrete numeric values. Analysis programs consider them to be measured on a nominal scale.
- 4) CHAR(string_length) specifies that a data item has as its value a string of variable length \leq "string length".
- 5) DATE data items store dates in an internal form. Dates are entered in a standard format, DDMONYY. DD=two digits specifying the day, MON=first three letters of the month (irrespective of capitalization), and YY=last two digits of the year. There are no spaces. The standard form is auto matically encoded into an internal "arithmetic date" for storage and numerical manipulation. Stored dates are converted back to external form by TOD retrieval modules.
- 6) CODE data items are stored as numeric values related by encoding-decodi procedures to a more legible representation. For example, the item "sex" might be coded as 0, 1 for F, M.
- 7) CONFIDENTIAL data items are encoded and decoded by private procedures providing keyword-protected scrambling. Only READER information can be confidential.
- 8) POINTER data items store pointers to information contained in an auxilia data file, defined for a specific TOD databank. What gets pointed to by a POINTER items is determined when an individual TOD databank is defined. As an example, a HEADER element of type POINTER might point to the first of a group of textfile lines comprising a reference letter for each patient.

There are three data type qualifiers. At data entry time an item for which LIMIT(min,max) is specified will have its value checked. If the entered value falls outside of the specified limits, an error message will be given.

At data entry time an item for which CHECK is specified will have its value checked for validity ty the user-provided encoding-decoding data check procedures / ,. . . [Se< iiL:- -I_ ii- .- L-~a~.j. iC;..,.5 C'S,' e*ei*iL'Zl Tir. 15 ~~e~~fie~ ;;ill PA--ye their values checked by a big data checking program run incrementally against the databank asynchronously with data entry.

The fifth attribute of a data item specifies the retrieval type for information stored in this item. Retrieval files will be maintained by a public UPDATE program facilitating the retrieval of data items in accordance with their retrieval types. Like the other attributes, retrieval type cannot be omitted from the formal assignment statement. Any subset of the types INDEX, RAmGE, TRANSPOSED, or PRIVATE may be given for retrieval type, or else the user must specify the type NONE. PRIVATE indicates that a user maintains private files to expedite retrieval for this parameter.

Leaving several formal array elements undssigned with attributes greatly decreases the cost of adding new data items at databank recompilation time, and is highly recommended. Assignment of previously unassigned formal array elements is a relatively inexpensive way of expanding a data bank, whereas redeclaration of formal arrays leads to costly reformatting of the entire databank.

The INITIAL statement causes the data items corresponding to a named formal array element to be set to an initial value given as "initialvalue" prior to data entry, checking and encoding. If "SAME" is specified for a HEADER element, it has no effect. If "SAME" is specified for a PARAMETER element, then it has the following effect: On the first visit for each patient, this parameter element is set to UNDEFII!ED, or to its initial value if VALUE (initial-value) was indicated. Then for each successive visit this item's value is initialized to its value on the previous visit for this patient. If no initialization is specified, numeric values are automatically set to "UNDEFINED" = -0.0 and string values are set to null as a default.

Formal array elements must be assigned attributes before they can be assigned initial values.

Declarations of formal arrays must precede use of formal array elements, and an assignment of attributes to a formal array element must precede its initialization. If several assignments or initializations are specified for one formal array element, only the last will be effective.

E. Required Items:

The following three HEADER and two PARAMETER element assignments must be included i. every databank definition if a user wishes to take advantage of public data entry and retrieval programs. The assignments are written in DDL with explanatory comments. Data type qualifiers and retrieval types other than NONE are acceptable throughout.

```
\texttt{HEADER}(\texttt{l}) = (\texttt{'Name})HEADER(2) 
                      {\tt last, first, middle\ initial', 'none', NAME, CHAR(30)},= ('Stanford medical record number','none',MEDREC,VALUE,NONE
            /* Stored as a six digit number. */ 
\texttt{HEADER(3)} = ('Birthdate','none',BIRTHDAT,DATE,
PARAMETER(1) = ('Date of visit','none',DATE,DATE,TRANSPOS
PARAMETER(2) = ('Age of patient to date','years',AGE,VALUE,NC
                 /* The value of age will be automatically computed from 
                   BIRTHDAT as a decimal number of years when the value 
                   of DATE is entered. */
```
F. Restrictions on Databank Schemas:

See ACME Note TDPT, Section 6.

 $-53 -$

T-TuudoL

Analyzing TOD Operational Costs

Farly during the TOD project we found that physicians did not have the tools available to get an adequate picture of the operational costs of a databank. They knew their total costs and what they did, but the task of allocating costs to individual transactions had become complex. An attempt was made during the project to set up a cost analysis framework to aid the researcher in determining where his computer dollar went. ACME Note TODCST outlines a study underway; ACME Note TODPDA outlines the method of trapping basic cost and transaction data for the TOD system. Note TODCST follows.

Analyzing the Costs of Running a TOD Databank

TODCST-1 Frank German0 March 22, 1973

This note outlines a study to analyze the costs of running a TOD databank. Components of costs and procedures to compare them are identified, both for comparisons of cost components within individual databanks and among several databanks.

The outline described in this note will be used to compare TOD costs in Immunology and Oncology, two operational TOD databanks. In the future, as the number of TOD databanks grows, the benefits of cross-databank cost comparison, as outlined here, will increase. In any event, cost comparisons among components of a single databank are always useful.

Raw Data Sources:

The raw data for analyzing costs of a TOD databank can come from several sources:

- (1) ACME end-of-billing period accounting detail which shows LOWN-LOGOFF pageminute charges by account.
- (2) TOD program TD-CSTCK output.
- (3) Independent logs kept by TOD users.

Each source yields information on some aspects of the cost picture. All have certain biases which should be understood.

ACME End-of-Billing Period Accounting Data:

The ACME end-of-billing period accounting data is the most critical because the monthly bill is developed from it. This is the number we are trying to analyze and justify. At present, an ACME user only gets a monthly bill showing totals for disk blocks, pageminutes, and terminal connect charges for the month. The logon-logoff detail is only provided on special request.

Since this detail is only identified by date, time, terminal, and pageminutes used during a session, and since a user can run many different programs during the time he is logged on, it is difficult to break down the costs of individual sessions to component operations. A component operation can be a program or some transaction(s) within a program.

The availability in ACME of a "pageminute used so far" function would be helpful in breaking down session costs. Even with this function, the responsibility of trapping the data and analyzing it falls to the application program. The availability of this command would eleminate the major error in accounting estimates made by the TD CSTCK program described below.

TOD TD_CSTCK Program:

This program was designed to analyze the data generated by the TODOPEN-TODCLOSE procedures present in mcst TOD programs. These procedures estimated the pageminutes used by individual TOD programs and logged the appropriate data. The pageminute function described above would turn these estimates into actual data.

The bulk of TD-CSTCK is a TOD cost analysis system. Attempts are made to allocate costs by program, but more important, by transactions done by the programs. See ACME Note TODPDA for a discussion of the TD_CSTCK program.

TD-CSTCK is the first step in the data analysis because it is designed to allow the user to purge the data once a month coinciding with the accounting cycle. Data over time, for which an analysis framework will be developed later in this note, is not presently kept in computer form. Later as the kinds of analysis become more stable, TD CSTCK can be modified to keep data month-to-month and appropriately analyze it.

Presently there exists a program TD COMPR, which is a modified TD CSTCK. TD COMPR will summarize costs for any TOD databank whereas TD-CSTCK can only access data from the name and project with which a user logged onto ACME. TD COMPR is only to be used by the TOD Manager.

Individual Logs:

Individual logs suffer from several disadvantages. They require everyone who operates the terminal to use it, a requirement which in practive is never adhered to. To get the detail provided by the method above would require an amount of time most users of the system would not be willing to spend. Individual logs are useful to keep data unrelated to the programs and transactions. For example, keeping a log of the names of people using a databank would allow better distribution of documentation to the people who need it.

COST Components for Individual TOD Databanks:

There are three major areas of cost in the operation of a TOD Databank. They are: personnel costs, computer run charges (including terminal rental), and computer disk storage. Personnel costs will not be directly studied here.

Computer Run Charges:

Computer run charges are composed of many components:

- *The terminal rental per month is presently \$190 per month.
- The actual program charges for running a TOD databank can be broken down into several functional area: data entry and update; subset; analysis; and utility procedures. Many TOD programs can be found in each area. ACME Note TODI indicates this partitioning.
- *The bulk of the computer run costs for a MD databank fall into the data entry and update area. Dr. Jim Fries has estimated data entry and update charges to be 75% of the run time costs.

Individual Aggregate Analysis:

A useful analysis for an individual databank would be to follow the individual group mentioned above over time. Graph I illustrates this analysis. "Costs" can be measured in dollars, pageminutes, or time-units.

Graph I. Individual Aggregate Analysis

Time, Months

Note: Figure not necessarily to 'scale.

Graph I illustrates an aggregate analysis and includes all the factors that influence costs, such as more people using the databank, more patients in the databank, higher activity than normal in a functional area. etc.

In order to begin to understand the reasons why these aggregates change, a more detailed analysis is required. Data entry and update costs will be used as an illustration because they are the most critical. A similar analysis could be developed for the other functional areas.

In TOD, a patient's demographic information can be entered; this data can be updated; a patient-visit can be entered; and this visit data can be updated. These operations are done by the four programs TD_ESTAB, TD_EFIX, TD_VISIT, TD_VFIX, respectively. Costs in each of these programs are influenced by the same variables. The major variables are: the number of patients or visits entered or fixed at one use of the program; the number of elements per patient which were entered or fixed; and the complexity of the data entered. These variables influence computer charges for data entry. They also influence disk charges because the files increase in size as more data is entered.

In looking at the costs of using these programs, the example of entering a patient's demographic information using program TD_ESTAB will be studied. Similar statements can be made for the other programs. In TD ESTAB, the cost (time, dollars, or pageminutes) per patient entered is useful because it re moves the direct effect of the number of patients entered to yield a more basic cost coefficient. For the same reason, the cost per patient element entered is taken as the elementary cost unit. One could argue that the complexity of the element measured in key-strokes to enter would be the true basic cost unit. At present we are only analyzing down to the element level.

Watching the cost per patient and the cost per element provides useful insight into the data entry costs. If the number of elements added per patient was always about the same, then we wouldn't have to use the cost per element; the cost per patient would do, In Oncology and Immunology, the number of elements entered per patient does vary widely. TD CSTCK does not presently include compile time.

One must remember that the cost of compiling a program is finite and could shadow the effect of entering data. The effects are relative in the sense that 20 elements for a patient might be comparable to one quarter of compile time. In this case, ignoring compile time, which is not accounted for by TD-CSTCK, introduces a sizable error, If 200 elements were entered, the error of not counting compile time becomes negligible.

Having picked cost per element as the elementary item, Graph II illustrates a method of studying it over time. Cost will be measured in minutes per patient element.

Graph II. Elementary Cost Components Vs. Time

One would expect variability in this elementary item for several reasons. A badly written TOMR chart with items improperly marked will slow down the data entry operation. Interruptions of the data entry person such as phone calls, visitors, etc., will increase the data entry cost. A new data entry person will take more time until he or she has learned the system. Assuming a consistent chart, no interruptions, and the same data entry person, the costs still will vary (e.g. data entry person can become bored or fatigued).

Cost Comparisons Among Databanks:

Aside from intra-databank comparisons discussed earlier, inter-databank comparisons can be made.

Studying the costs of running databanks using the same programs allows us to establish norms or standards for the operations involved in running a databank. The data is also useful in estimating future costs of planned databanks. A better understanding of costs will allow the detection of high cost areas and, hopefully, the subsequent improvement of these areas. For example, if a databank has an elementary cost component twice that of another databank, that tells us something about the first databank's operation. Moreover, if we can determine what the second databank is "doing right" and can teach the first databank the same procedures, it too can lower its costs.

Each month, elementary cost items can be tabulated and summarized for all the operational TGD databanks. Graph III illustrates a data entry cost comparison for several users.

Aggregate items can also be compared among TOD users. Some of the more important items are listed in Table I. All items are per month.

Table I. TOD Costs Comparison for all TOD Users

Month of: xxxxxxx

ITEM USER 1 USER 2 USER 3 ...

From ACME Accounting:

Terminal Cost. Computer Cost. Number of Pageminutes. Number of Minutes. Number of Blocks.

From TD-CSTCK:

Minutes. \$ Estimated. Pageminutes Estimated. Cost per Program. Cost per Functional area.

Miscellaneous:

Ratio($$ from ACME to $ from TD CSTCK$). Ratio(Time from ACME to time from TD CSTCK). Ratio(Pageminutes from ACME to page. from TD_CSTCK). % of cost falling in each functional area. Number of patients in file. Number of patients entered, Number of patient-visits in file. Number of patient-visits entered.

A final analysis which would illustrate the use of TOD is shown in Graph IV.

March 1973

Time, Months

$-67-$

Dist: Staff/TOD/All

۰.

C. New Analysis Programs for Time Series Data Project: ACME

Realtime

Investigator: Will Gersch

Dept. of Neurology

1. Description

ARSPEC is an interactive research tool for automatic spectral density analysis of time series data. The program performs four interrelated tasks: (1) display of time series data; (2) filter design and application to data; (3) data reduction; and (4) spectral density computation. The user instruct ARSPEC to perform tasks in a desired order by issuing supervisory commands.

Unlike PUBLIC program TIMESER, ARSPEC uses an automatic decision procedure to produce accurate spectral estimates. This technique is applicable to all time series data.

The computational procedures are defined in ACME Note EARSPE-2.

2. Historical Note

The recursive procedure for computing autoregressive coefficients is due to Levinson $(A)(1947)$. This procedure was "re-discovered" by Durbin $(B)(1960)$. The automatic decision procedure for deciding the order of the autoregressive model is due to Akaike (C,E)(1970). An alternate statistical procedure based on earlier statistical procedures for deciding autoregressive model order appears in Gersch (D)(1970).

3. A Short Explanation of Spectral Analysis by Autoregressive Model Technique

Conventional spectral analysis procedures compute spectral density estimates using a weighted Fourier transform of the empirical autocorrelation function of observed time series. The use of conventional techniques is complicated by the requirement that a user assign the values of parameters determining statistical tradeoffs in the spectral representation. This requires expertise with spectral estimation theory. As a case in point, conventional spectral analysis techniques are employed by ACME PUBLIC program TIMESER.

The autoregressive model technique fits a model to be observed data. This model is autoregressive in that each observation of the time series is expressed as a linear combination of its own past (hence it is regressive on itself) plus a term drawn from an uncorrelated sequence (Equation 1). The coefficients of the model are determined by a least squares fit to the empirical autocorrelation function. The order of the model is determined by the Akaike (E) final predictor error criterion.
Core Research & Development (Continued)

Spectral density estimates are calculated from the coefficients of the autoregressive model (Equation 2).

The statistical performance of spectral estimation using the autoregressive technique has the properties that (i) the spectral estimates are unbiased, and (ii) the variance of the spectral estimates is approximately given by

$$
\text{var } \hat{S}(f) = \frac{2p}{N} \hat{S}^2(f)
$$

where $p =$ the order of the autoregressive model. The statistical performance of this procedure is at least as good as the performance of the best conventional spectral estimate. Finally, the fact that the procedure is unbiased eliminates the need for expert determination of the statistical trade-offs, and hence the procedure can be made automatic.

- 4. References
	- A. Levinson, N. (1947). The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction. 5. Math. Physics, z, pp. 261-278.
	- B. Durbin, J. (1960). The fitting of time-series models. Rev. Int. Stat. Inst. 28, pp. 233-244.
	- C. Akaike, H., Information theory and an extension of the maximum likelihood principle, presented at the Second International Symposium on Information Theory, Tsahkadsor, Armenian SSR, 2-8 of September, 1971. (To be published in Problems of Control and Information Theory, USSR 1972.)
	- D. Gersch, W. (1970), "Spectral Analysis of EEG's by Autoregressive Decomposition of Time Series", Mathematical Biosciences, 7, 1970, pp. 205-222.
	- E. Akaike, H., Statistical Predictor Identification, Ann. Inst. Stat. Math., 21, 1970.

5. Sample Runs

Raw data is a sine wave imbedded in noise. Noise is Gaussian noise with mean zero and standard deviation an order of magnitude greater than the sine wave amplitude. Noise has been passed through a high pass filter before addition to the sine wave.

Core Research & Development (Continued)

 \sim \sim

Data set number of graphics device=?13 UJ YOU ARE NOW PERMITTED 01 LINES ON TVE 1800 Number of raw data samples=?1000 Name of raw data file =?TEST Key of record containing raw data=?1 Enter a command =?display Enter 0 to operate on raw data, nonzero to operate on processed data=?0

Enter a command =?spectrum

Enter 9 to operate on raw data, nonzero to operate on processed data=?0 \equiv 'Autoregressive model is of order',= $49:$ Optional description of spectrum =?sine wave in filtered noise Enter 0 if done, nonzero to save spectrum on data file.=?0

> SINE WAVE IN FILTERED NOISE Ø -12 \mathbf{B} POWER SPECTRUM IN -20 $-3x$ $-4₁$ \overline{c} .0 0.1 0.2 α . 3 2.4 0.5 NORMALIZED FREQUENCY

Enter a command =?filter High frequency cutoff point=?0.1 Low frequency cutoff=?0.05 Enter 0 for low pass, nonzero for high pass filter=?0 Enter 0 for unsmoothed, nonzero for smoothed filter=?1 Enter 0 to display, nonzero to apply filter=?0 Enter 0 to operate on raw data, nonzero to operate on processed data=?0 FILTER DESIGN BO \overline{L} TRUE TRANSFER FUNCTION $-10.$ -22 $-30.$ -40 າ
0.5 $\overline{0.3}$ \mathbf{a} .4 0.1 0.2 b.o NORMALIZED FREQ.

 $\frac{1}{1}$

After filtering, we can see the original sine wave. Enter a command =?display Enter θ to operate on raw data, nonzero to operate on processed data=?1

PROCESSED DATA

FIGURE 4

Note the end effects caused by filtering

2. The second example is a spectral analysis of actual EEG data.

Pata set number of graphics device=?91 Humber of raw data samples=?2000 Hame of raw data file =?EEGDATA Key of record containing raw data=?100 Enter a command =?display Enter 0 to onerate on raw data, nonzero to operate on processed data=?0

OFILE: EEGDATA, DREC: 100

Enter a command =?spectrum Enter 0 to operate on raw data, nonzero to operate on processed data=?0
='Autoregressive model is of order',= 22; ='Autoregressive model is of order', = Optional description of spectrum =?EEG SPECTRUM Enter 0 if done, nonzero to save spectrum on data fIle.=?O

A low pass filter is now applied; then data is reduced to 1000 points.

```
Enter a command =?fi 1 ter 
High frequency cutoff point=?0.25
Low frequency cutoff=?0.2
Enter 0 for low pass, nonzero for high pass filter=?0
Enter 0 for unsmoothed, nonzero for smoothed filter=?1
Enter 0 to display, nonzero to apply filter=?1
Enter 0 to operate on raw data, nonzero to operate on processed data=?0 
Enter a command =?reduce 
Enter 0 to operate on raw data, nonzero to operate on processed data=?1
Number of reduced data samples=?1000
Cnter a command =?spectrum 
Enter 0 to operate on rat:4 data, nonzero to operate on processed data=?1 
 \blacksquare'Autoregressive model is of order',\blacksquareOptional description of spectrum =? 
Enter 0 if done, nonzero to save spectrum on data file.=?0
Enter a command =?done 
0 127: RUN STOPPED
   AT LINE 3.100 IN PROCEDURE ARSPEC
```


The spectral outline of the reduced and filtered data has characteristics parallel to the original data. Note that 50% reduction of data causes the normalized frequency scale to double.

 \overrightarrow{c}

 $\begin{array}{c}\n\text{char}\rightarrow \\
\text{char}\rightarrow\n\end{array}$ $\mathbf{1}$ <optional description of spectrum> Format for displaying filter and spectrum (tcharx, tchary)=(center, 935) *gridheight=800 * xfmt = image statement containing Fortran format for x axis labels y fmt = image statement containing Fortran format for y axis labels horizontally displaced 2 rasters ी
पुर knumeric vertically centered label> VICKENDIX- \mathbf{I} \vec{a} *gridleft=150 ω tic displaced by 2 rasters ycharx, 2 *tic $\xleftarrow{-} \rightarrow$ charh vertically & horizontally Ö. $\overline{8}$ tic
S ychary) = (35, 200) Λ N <numeric label> $\ddot{\bullet}$ 510 *gridlength=800 $\mathbf{\hat{v}}$ *charh=30 $200t = 100$ *These are the actual program names of

 $\pi^*(xchary, xchary) = (300, 0)$

 \mathbf{v}

parameters defining the output format, specified with the number of rasters they are initialized to. To modify these settings, change the first declaration at the beginning of ARSPEC.

Core Research & Development (Continued)

D. Miscellaneous Core Projects

l.

For descriptions of the other projects which had core research status during FY73, see the Summary of Computer Resource Usage, Categories 5 and 6, in Section VII of this report.

VI. BUDGET

The budget section of this annual report is divided into the following topics:

- A. Resource Expenditures
- B. Expenditure Details
- C. Summary of Resource Funding
- D. Resource Equipment List
- E. Income Statement for 12-Month Period Ending April 1973

Fiscal year 1973 direct costs will be \$800,000. This will be offset by approximately \$260,000 in income from service fees. These figures exclude costs associated with providing terminal services to users and the income received through terminal service fees.

Section E indicates that income of \$289,000 was received from users in the 12-month period ending April 1973, compared to prior year income of \$230,000. The terminal service fees accounted for approximately \$100,000 in additional income during each of these two years; roughly one-half of this sum was spent for terminal rentals while the balance has been used to provide services best offered from a central point and to support new communications gear for use of the 370/158 facility.

A. Resource Expenditures

SUMMARY

 $\mathcal{A}^{\mathcal{A}}$

Includes education courses (1)

 (2) Assumes \$403,561 exempt equipment costs
and user income of \$260,000

* 06S refers to one year extension

B. Expenditure Details

DIRECT COSTS ONLY

 ζ

B. Expenditure Details

DIRECT COSTS ONLY

August 1, 1971- August 1, 1972-
July 31, 1972 - July 31, 1973

3. Equipment

 \mathcal{A}

 $\mathcal{L}^{\mathcal{L}}$

Major Equipment

* Reflects corrections of charges mistakingly placed
on Grant account for graphics communication services.

B. Expenditure Details

DIRECT COSTS ONLY

Purchased Equipment

* Some money has been shifted from other budget categories
to permit purchase of equipment needed to provide follow-
on PL/ACME services. More detail is presented on separate
correspondence to NIH.

B. Expenditure Details
DIRECT COSTS ONLY

 $\label{eq:2.1} \frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\right)^2\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\right)^2\frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\right)^2.$

 $\frac{1}{2}$

 $- 84 -$

C. Summary of Resource Funding

BUDGET PERIODS

* A request for additional funding support has been forwarded to NIH.

D. Resource Equipment List

RENTAL EQUIPMENT

 $\sim 10^7$

360/50 Configuration

 \mathbf{F}

 ∞

 $\bar{\Gamma}$

D. Resource Equipment List

RENTAL EQUIPMENT (Cont.)

 \sim

Supporting Equipment Rentals

 \mathbf{I} 25 \mathbf{L}

Contract Contract

Contractor

 $\sim 3\%$

PURCHASED EQUIPMENT

Period Covered -- 8/1/67-4/30/73

VII. UTILIZATION DATA

A. Interpreting Utilization Charts.

The terms used to discuss ACME utilization involve charging units and categories of users.

1. Charging Units

Last year, the computer service charge units were:

pageminutes terminal connect time blocks of disk storage terminal service charge

A pageminute is defined as occupancy of 4096 bytes of core for one minute. Terminal connect time is the total number of minutes that a terminal is connected to the system in a logged-on condition. A block of disk storage is a fixed-length block of 2000 bytes of 2314 -type disk storage. The terminal service charge covers monthly terminal rent plus other services offered by the ACME staff. This service charge is handled by the University independent of the ACME grant.

2. User Categories

The table below shows the category identifier, definition, and rate of each user category. The rate charged per pageminute varies by user category, and some categories are subsidized 100% by the ACME grant. An asterisk next to the category identifier designates those so subsidized. All other categories are paying. There is a distinction between realtime and non-realtime users. Realtime users use the 1800 processor or 2701 data adapter for data collection or process control functions.

PAGE-MINUTE CHARGE TABLE

Category cents/pageminute

Page-Minute Charge Table (Continued)

*No cash charges, i.e., absorbed by the ACME project budget.

B. Patterns of Use.

are in execution at any point in time. This is consistent with a leve Approximately $2/3$ to $3/4$ of the terminals logged-on to the system achieved one year ago. Somewhat fewer terminals are logged-on each day in recent months. This reflects the fact that more people are being charged for their utilization. It also reflects the fact that fewer new users have been recruited this year due to the uncertainty of PL/ ACME's future. The number of terminal hours did not decrease significantly despite the increase in the number of 30-character-per-second terminals. Subjectively, it appears to the facility director that much of the utilization today is based upon programs written during the past three or four years. Perhaps as many as 50% of the users rarely write new programs for data entry and data analysis.

During the past year approximately 85% to 90% of the available disk storage for users has been used. Frequently we have run out of space on individual disk packs during the normal operating hours. This has caused considerable inconvenience to all users. One additional disk drive was added to the system this year to alleviate the problem.

'The Utilization by Department table indicates considerable usage by the Computer Science Department. The DENDRAL project accounts for essentially all of this usage.

The final table in this section presents a summary of all utilization by all users during the year. On this table, the designation "C" (collaborative) indicates that the project (1) received programming services greater than normal consulting, (2) was accommodated by special changes to the system by the ACME staff, or (3) was involved in a normal research collaboration relationship with ACME.

 \overline{f}

 \sim

 $\sim 10^{-1}$

Twelve-Month Period Hay 1972 - April 1973

*Distribution is that of April 1973.

 \mathbf{u}

 $\frac{1}{2}$.

o.

Twelve-Month 5/70-4/73 Page 2

Department/Division	# of Terminals		PAGEMINUTES students non-chargeable	chargeable		BLOCKS students non-chargeable chargeable		students	TOTAL CHARGE non-chargeable chargeable	
HOSPITAL										
Cardiology	1.00									
Clin Lab - Immunology				6,219			72			114.16 Æ.
Clin Lab - Infec Diseases	1.00			1, 119, 523			88, 324			24, 323.66
Clin Lab - Pathology	1.00			511, 269			7,404			5,787.34
Data Processing	1.00			501, 615			32,962			9,046.46
Pharmacy				8,006,546			95,581			26,957,57
HOCPITAL TOTAL	4.00			10, 145, 172			224, 343			\$66,029.19
CLINICS	1.00			515,400			23, 350			\$7,828.71
OTHER CLINICS AND HOSP	1.00			680, 758			18, 919			15, 874.73
CAMPUS										
Aero and Astro				$-0-$			624			\$ 62.40
Biosciences	2.00	3,035		286,747	135		6, 343	\$76.15		3,617.73
Chemistry	1.00		176, 223	199.075		15,697	14,478		\$4,846.90	4, 827.51
Computer Science**	3.00		4,597,329	1,606,744		176, 473	62, 854		41, 210, 20	20, 025.17
Lav				$-0-$			696			60,60
Math	1,00			5,267			70			98.36
Physics				7,723			2,954			4.28,00
Psychology		867		28,093	51		978	22.81		662,08
Statistics	.50			5,886			363			148.20
CAMPUS TOTAL	7.50	3,902	4,773,552	2, 139, 515	186	192, 170	89, 360	98.96 \$	\$46,057.10 \$29,939.05	
UNKNOWN SCRATCH ***		216, 629			161			\$4,450.82		
OTHER		<u>20, 736</u>	3,139	474,996	126	53	24, 992	\$436.24		69.41 \$11, 396.27
USER TOTALS	41.00	833,046		6,756,805 21,522,568	38, 372	278, 913		895,701 \$20,888.98	\$86,776.57 \$289,084.12	
ACME	10.00		8,928,799****			198, 823			\$105,581.53	
GRAND TOTALS	51.00	833,046		15, 685, 604 21, 522, 568	38, 572	477,736		895,701 \$20,888,98	\$192, 358.10 \$289, 084.12	

 $\sim 10^{-1}$

 $\ddot{}$

 $\ddot{}$

*Distribution is that of April 1973.

"List include DENDRAL project, serving Genetics, Chemistry, and Computer Science Departments.
**Primarily the DENDRAL project, serving Genetics, Chemistry, and Computer Science Departments.
***Unknown users, mostly medical

 \cdot

 $\Delta \sim 10^{11}$ km s $^{-1}$

GRAND TOTAL PAGEMINUTES: $36,041,218$
GRAND TOTAL PLOCKS: $1,411,809$

 $\ddot{}$

April 26, 1975

and the control of the con-

 \mathcal{L}_{max} and \mathcal{L}_{max} . The \mathcal{L}_{max}

Numher of laer Projects By Department and Category at April 16,

 ~ 100

 \sim t .

 $\sim 10^7$

 \sim $\ddot{}$

 $\sim 10^{-1}$

I

n.

 $\tilde{}$

 \sim

the first state of the first state of the state of the state

 $\sim 10^{-1}$ km $^{-1}$ km

 $\left\langle \frac{1}{2} \right\rangle = \left\langle 1 \right\rangle$, $\left\langle \frac{1}{2} \right\rangle = \left\langle \frac{1}{2} \right\rangle$, $\left\langle \frac{1}{2} \right\rangle = \left\langle \frac{1}{2} \right\rangle$

 \sim

I $\frac{34}{4}$ I

ACME INCOME

 \mathcal{L}

 \sim

- g6 -

 $\bar{\pmb{\epsilon}}$

 $\overline{}$

 \mathbf{r}

 $\bar{\Gamma}$ $\bar{\Gamma}$ $\bar{1}$

 χ

 $\frac{1}{4}$

ACME INCOME

FY69 through FY73

*User charges instituted in March of 1969.

ACME UTILIZATION DATA

FY7Q - FY73

As measured by number of accounts (staff and operations accounts excluded):

As measured by income:

'Income provides for terminal rental, services best provided by a _rcentral facility, and some recharges for engineering services. ζ Rate reduction introduced in February 1970. Rate increase introduced in May 1972.

As measured by timesharing units:

d One pageminute represents occupancy of 4k bytes of memory for one minute.

100 \mathbf{I}

 \mathbf{r}

BURGHARY OF CONFITER RESOURCE HOAGE $\sqrt{2}$ April 1/, 1//2 - April 10, 1//3

 $*Cor = Core Research and Development
\n $C = Collaborative$$

 $S = Service$ $T = Triating$

 $\label{eq:2.1} \mathcal{L}(\mathcal{L}^{\text{max}}_{\mathcal{L}}(\mathcal{L}^{\text{max}}_{\mathcal{L}}),\mathcal{L}^{\text{max}}_{\mathcal{L}}(\mathcal{L}^{\text{max}}_{\mathcal{L}}))$

 $\bar{\beta}$

SUMMARY OF COMPUTER RESOURCE USAGE

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

C = Collaborative

S = Service

T = Training

-
-

SUMMARY OF COMPUTER RESOURCE USAGE

 \sim σ $<$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

 $C = Collaborative$

 $S = Service$

April 17, 1972 - *Ppril 16*, 1973

 $\frac{1}{\sqrt{C}}$ = Core Research and Development
 $C = \text{Collaborative}$
 $S = \text{Service}$
 $T = \text{Training}$

 $\sim 10^{-1}$

 \mathcal{L}

 $\ddot{}$

 $-104 -$
April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

C = Collaborative

S = Service

T = Training

 $\sim 10^7$

April 17, $1972 - \text{April } 16$, 1973

*Cor = Core Research and Development
 $C = \text{Colinheartive}$

S = Service

T = Training

- -

 \mathbf{L} 106 \mathbf{L}

SUMMARY OF COMBUTER RESOURCE USAGE April 17, 177° - April 16, 177°

*Cor = Core Research and Development
 $C = \text{coilaborat two}$

S = Service

T = Training

 \sim

 k_{p} ril 17, 1772 - April 16, 1773.

*Cor = Core Research and Development
 $C = \text{Colleventive}$

S = Service

-
- $T = Triating$

SUMMARY OF COMPUTER RESOURCE USAGE $Apr11.17$, $1972 - 4Apr11.16$, 1973 .

*Cor = Core Research and Development
 $C = C \cdot 1$ aboutative

 $S = Service$

 $T =$ Training

and the control of the control of

 \sim

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development
 $C = \text{Collaborative}$
 $S = \text{Service}$
 $T = \text{Training}$

 \sim

SUMMARY OF COMFUTER RESOURCE USAGE April 17, 1972 - April 16, 1973

 \sim χ .

*Cor = Core Research and Development
C = Collaborative

 $T = \text{Training}$

 $S = Service$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

 $C = \text{Collaborative}$
 $S = \text{Service}$
 $T = \text{Training}$

 \mathbf{L} 112 \mathbf{I}

April 17, 1972 - April 16, 1973

 $\begin{minipage}{.4\linewidth} \textbf{*Cor} = \texttt{Core Research} \textbf{and Development} \\ \texttt{C} = \texttt{Collaborative} \end{minipage}$

 $S = Service$
 $T = Training$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development
C = Collaborative
S = Service
T = Training

-
-
-

 $\mathcal{L}_{\mathbf{r}}$

April 17, 1972 - April 16, 1975

 $\sim 10^6$

*Cor = Core Research and Development
C = Collaborative

 $S = Service$
T = Training

 $\bar{\beta}$

April 17, $1972 -$ April 16, 1973

*Cor = Core Research and Development
 $C = \text{Cell}$ aborative
 $S = \text{Service}$
 $T = \text{Training}$

 \mathbf{I} 116 \mathbf{I}

April 17, 1972 - April 16, 1973

"Cor = Core Research and Development

 \mathbf{I}_c 11_L \mathbf{L}

 $C = Collobarative$
 $S = Service$
 $T = Training$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development
 $C = \text{Collaborative}$
 $S = \text{Service}$
 $T = \text{Training}$

 $\sim 10^{-1}$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development
C = Collaborative
S = Service
T = Training

 $\mathcal{A}^{\mathcal{A}}$

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

C = Collaborative

S = Service

T = Training

120

SUMMARY C. LOWFUTER RESOURCE USAGE

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

C = Collaborative

S = Service

T = Training

April 17, 1972 - April 16, 1973

*Cor = Core Research and Development

 Γ **122** \mathbf{I}

 $\bar{1}$

 $C = \text{Collaborative}$
 $S = \text{Service}$
 $T = \text{Training}$

RESOURCE USAGE

pril 16, 1973

*Cor = Core Research and Development
C = Collaborative
S = Service
T = Training

-
-

SUMMARY OF COMPUT

April 17, 1972 -

due to failures

Percentage of scheduled time

 $\ddot{}$

 $\langle \hat{f} \rangle$

not available to users

 $2.1%$

VIII. APPENDICES

- A. Rationale of Shared Data Base Concept
- B. Medical Center Computer Planning Chronology
- C. Small Machine Multiplex (Excerpt of ACME Note HAD.)
- D. ACME Note Index (May 11, 1973)
- E. ACME User Publications

 $-125 -$ APFENDIN A

DATE: March 16, 1973

To : P. Carpenter

From : E. Levinthal

SUBJECT: Rationale of Shared Data Base Concept

This memorandum elaborates the rationale supporting the shared data base concept as it relates to the Medical Center Computing Facility. Computing at the Medical School can be categorized as follows:

- 1. Use of in-patient or out-patient data by clinical departments. This involves, in varying amounts, three components, a) teaching, b) research, c) patient-service management (i.e. fees, records, bills, etc.).
- 2. Non-patient related computing by clinical departments. This part is only indirectly related to the shared data base issue. Insofar as clinical faculty are using a computer resource for their patient related computing needs, they are apt, as a matter of convenience and familiarity, to want to use the same facility for the remainder of their needs. This will be modulated by considerations of price and services offered.
- 3. Non-patient related computing carried out by non-clinical faculty. This is clearly unrelated to the data base concept. It is, of course, related to the cost and services offered on the 370/158 system compared to those offered elsewhere. Many of these users take advantage of functions which call statistical routines and which are now built into the PL/ACME system. These will also be an important requirement of users in categories 1 and 2. In this case therefore the issue (as in category 2) is the service rendered, not the data base.

Addressing solely category one, memoranda were solicited from several members of this class. The responses are attached^{*} and provide support for the shared data base concept.

There is clearly a momentum to use computers to handle patient related problems. Faculty are able to find resources to pursue these problems and will pursue them whether or not a shared data base in a central computer system is available. In principle the communication link researcher-to-researcher and researcher-to-business office are transactions that can be carried out by movement of paper or digital tapes or hardware interfaces between stand-alone facilities.

'%Iemos attached from Drs. Cohen, Fries, Harrison and Merigan.

In practice linkages are only made when the perceived need is deemed worth the effort. Without a central system the "potential" barrier to forming linkages can involve costly software and hardware interfaces. In a clinical research and teaching environment the number of possibly useful combinatorial linkages is large. If the "potential" barrier is great, innovation and experimentation is impeded. The forces in the system are then centrifugal instead of centripetal.

Since the management of the clinics and hospital also depends more and more on computer manipulation and extraction of data, the total systems behavior will have important economic as well as academic consequences.

The proposal for Computer Health Care Application Research gives an insight into the clinical and academic benefits of a shared common data base. This and several other grant proposals, involving interdepartmental collaboration have called for linking of data bases. The second section of the proposal addresses the important problem of the definition of the data base. The third portion, deals with file and retrieval systems for a clinical data base. This involves the potential utilization by twelve specific clinical activities of a shared data base. This grant, if implemented would spend approximately \$86,000 per year in computer services. Some of these developments are currently underway on ACME.

In addition to the academic research needs of the clinical faculty, there are the requirements of the hospital for a shared data base. These are derived from managerial and economic imperatives as well as the hospital's educational and research goals.

There is no current completely acceptable solution that meets the requirements of a complete Hospital Information System (HIS). The search for this solution is a very important problem and one in which Stanford should be involved. It will affect many aspects of medical education and teaching as well as practice within a hospital environment. Within the next several years many elements of such systems will be successfully implemented and will be important parts of the operation of Stanford Hospital. The 370/158 has the capacity to allow Stanford to implement a hospital information system. The design of such a system and the timing and funding of its implementation are not part of this plan.

The Technicon HIS at El Camino provides insights into costs and CPU requirements of HIS. From the operation of the El Camino system since the first of this year, it now looks like they will in fact realize net savings of \$85,000 per month, most of which will be realized by reductions in nursing staff personnel. El Camino is a hospital with 446 beds and 60 bassinettes. The Technicon Hospital Information System is designed around two 370/155's to support 2,000 beds at \$6.00 per day. The CPU cost is about one-third the total cost. This is in addition to the cost of business operating systems. Roughly, this says that implementation of such a system at Stanford with 612 beds and 57 bassinettes would approximately double the dollars that would be available to be spent by the Hospital for central processing over our worst-case projections and a 50 per cent increase in our conservative projection in whatever year an HIS should be installed.

P. Carpenter March 16, 1973

It will be economically important in the future to bring together dispersed elements of a patient information system into a coherent whole. It may be too difficult and expensive to do so, if dispersion has gone on too long. This is the difference between a stand-alone community hospital and a hospital-cum-medical school. The former can wait until it knows exactly what it wants to do. Stanford Medical School faculty and their research and teaching interests are in integral part of Stanford Hospital. They will and should carry out their academic functions in the best way available to them. Nothing can or should stop the dispersive process except the better alternative of a well-managed reliable central system that by its very nature makes collaboration easy.

ECL/mla

Attach.

To 1 Elliott Levinthal, Ph.D.

FROM : Stan Cohen, M. D. Houle n. Please

SUBJECT: Need for Common Computer Facility at the Medical Center

As we have discussed previously, there is an important need for a computer facility at the medical center to provide capability for faculty to share programs and data related to both clinical activities and research projects. At the present time, individual projects being carried out by various faculty members constitute component parts of what will probably eventually develop into a hospital information system capable of handling large amounts of patient-related data. Included among these components are the drug interaction warning system of the Division of Clinical Pharmacology, the Microbiology laboratory system developed on ACME by Dr. Merigan and his collaborators, the Clinical Chemistry and Hematology laboratory system developed by Dr. Sussman, the Medical Records system of Dr. Jim Fries, and the Cardiology data system of Dr. D. Harrison,

Patient care at this medical center requires that these separate data bases be available on a central computer system so that information accumulated by one project can be shared by others. For example, the identity of organisms cultured by the Clinical Microbiology laboratory and their resistance pattern should be accessible by the pharmacy system programs, so that a prescription that is Inappropriate for a particular organism or drug resistance pattern can be detected at the time it is filled, Similarly, data being accumulated by the clinical chemistry laboratory indicating inadequate renal function should be available to the pharmacy system, so that alteration of dosage may be made for a drug eliminated from the body by excretion through the kidneys, Conversely, drugs that artifactually influence the results of laboratory test findings by interference with spectrephotometric determinations and other test procedures, and this information should be available to the clinical laboratory. Cardiology data should be available for similar reasons, 'and since drug influence interpretation of cardiovascular tests, pharmacy data should be available through the cardiology system. All of the types of data indicated here, plus clinical findings related to the patient history and physical examination should be part of the time-oriented medical record system being developed by Dr. Fries.

Although this brief memo strasses.the patient-care benefits that would derive from having a large medical center computer system available for sharing of data bases and programs, I also want to emphasize the importance of such a system t6 faculty research. Linkage of the clinical microbiology laboratory and pharmacy systems will enable epidemiological investigations of the effects of antibiotic use on resistance patterns of organisms isolated from patient populations. Similarly, research to detect new effects of drugs on clinical chemistry tests will also be feasible if the data bases can be shared. Atlhough these are just a few examples, there are many other instances where sharing of data bases will enable important investigative questions to be asked and answered.

I hope that this brief memo provides the information you are seeking.

- 129 -

DATE: March 8, 1973

To : Elliott Leventhal

floM , James F. Fries

SUBJECT: Advantages of a variety of medical database operations sharing the same computing equipment.

> Within the medical center and hospital there are a number of patient related computer databanks. Inevitably, the number and variety of clinical databank operations will increase over coming years. Material included in these databanks will be diverse yet similar. Thus, patient identifying infonnation, financial and accounting information, clinical information required for insurance and third party carriers, historical and physical examination data elucidated by physicians, therapy prescribed and drugs dispensed, and the multiple forms of information emanating from various clinical laboratories, x-ray, cardiac catheterization and pathology departments will be accumulated in computer databanks. Over the long term, the facility with which information may be exchanged between these different operations will be of great importance, A research study may require stratification in terms of socio-economic data kept by the business office. The business office may require clinical information available in other databanks to process insurance forms. Billing may ultimately be related to the actual provision of the service at the physician level as documented in the chart and from laboratory information as it becomes available to the physician. Without provision for linkage and exchange of information the individual databank operations will require duplication of effort in data entry. Without capability of linking laboratory computer systems to clinical medical record databanks, laboratory data must be manually re-entered.

It can be stated fairly that medical computing has consisted in large part of duplication of effort both at Stanford and elsewhere. As the need for computer based clinical information systems grows there is the possibility of ever greater fragmentation and duplication of effort. The existence of a central computing facility for the medical center and hospital will allow <code>planned</code> growth, <code>minimal</code> redundancy, and exchange and pooling of clinical states and $\boldsymbol{\theta}$ data. It will place the hospital and medical school in a strong positi to meet increasing governmental requirements for "quality assurance" and medical audit.

JFF/hcp

....... 13 iu*i*u

MAGE A المتكافئة

March 8, 1933

Dr. Elliott Lcvinthal Genetics Department

Dr. Donald C, Harrieon Cardiology Division

Advantages for a Hospital Computing System

Following our discussion yesterday, I have considered the advantages of a medical school computing-system which would be .a combination of hospital and medical school programs. The overall advantages are as follows:

- A. Having a joint facility in the medical center would permit a common data base for all patients. This is essential for ongoing clinical research and for ease in efficiency of administrative operations. The case for this 1s as follows:
	- 1. All patients under the care of Stanford faculty members in the Stanford University Hospital are either referred from the Stanford clinics prior to admission or are seen in follow-up in the clinics. Thus, it is essential that a data base include both aspects of the patient's record. This would encompass laboratory reports, x-ray studies, and ongoing follow-up data. These patients are frequently part of research protocols relating to the action of specific drugs, to the effects of surgical procedures, etc., and represent the basis for much of the clinical research being carried out.by the clinical faculty.
	- 2. A patient seen by one particular group in the hospital is frequently seen by others and data common to studies being carried out by several interrelated groups should be
available to the various division and departments. This available to the various division and departments. is particularly true in the case of Cardiology where .patients .are first seen by the medical cardiologist. Data are accumulated on the patients by the clinical laboratories, by the x-ray units, by the cardiologic units with special computer facilities such as the catheterization laboratory or the EKG laboratory, and then the patients undergo some surgical procedure in the Surgery Department. These patients are then followed up jointly by the various members of the Medicine, Surgery, and Radiology faculty. Consultants from Infectious Disease, from Immunology, and from other disciplines also frequently are asked to see these patients. To develop new concepts regarding the pathogencsis of disease, to test this in clinical populations, and to determine the effects of interventions upon these diseases, it is essential that theso groups interrelate their data.

Dr. Elliott Lovinthal -2- March 8, 1973

- 3. At the same time clinical data are being transmitted to patient's records, hospital charges can be assessed. Thus, for ease and efficiency of administrative detail, a cooperative computer facility is necessary.
- B. With increased emphasis upon judging the quality of medical care and upon determining cost effectiveness of care, the integration of hospital activities and medical school activities becomes absolutely essential. Computer surveillance for drug interactions, for physician performance, and for developing now educational activiticg related to this aspect of medicine, nccoaaitate a combined hospital medical school computer facility.
- c. The accumulation of a critical mass of individuals working in hospital information systems for Stanford Medical School seems essential. The interrelationships of data from small computer systems in the various divisions and departments and support for these interfaces would bo provided by a combined computer facility.

For the reasons of improving the delivery of health care, for enhancing clinical research, and for improving integrated teaching programs I would strongly support the developmont of a hospital medical school computer facility.

DCIIrgr

DATE: March 9, 1973 . . ,., MAR 1 3 1073

I

ء
3

č $\overline{\mathbf{c}}$

 \bullet

VELLER

うちさつえるこうころ

 \bullet

SIANTURU

From : Thomas C. Merigan, M.D. Chief, Division of Infectious Discasss

SUBJECT:

This memo is in respofise to your questions in regard to my thoughts concerning the ACNE system and its present and future contributions to clinical investigation. The availability of equipment and the ease of the language of ACNE has personally benefited me enormously during the past 5 years while we have been working with patient oriented systems forour Diagnostic Microbiology and hospital epidemiology functions. As you know, all'of our antimicrobial identification and antibiotic sensitivity data goes into ACNE on an on-line basis from our hospital service laboratory. This involves only a minimal amount of time for our secretaries and technicians and produces a useful return from t-o standpoints: the antimicrobial sensitivity data is quality controlled prior to its issuance to physicians and all of our previous experience is immediately accessible for our clinical consultants as well as the Diagnostic Microbiology Laboratory personnel.

In regard to hospital epldcmiology, the filed information is automatically put together on a monthly and semi-annual basis for reporting to the Infection Control Board members and the state and county authorities. The infection control nurses use this information in deciding whether there is any increased incidence of nosocomial infection at Stanford University llospital, and now records dating back two years are available in that area whereas the antibiotic sensitivity and isolation information goes back some four years allowing many types of comparisons which wouldn't be possible without this regular recording of data.

I think the point you are particularly interested in, however, is how a commonly shared system among various clinical users which is tied in with the hospital system might be particularly advantageous. We find that as the ACNE system was used for development and now the maintenance of our infection control and diagnostic microbiology systems, these two systmes can be linked up quite easily and personnel who operate one can also utilize the other. However, a very exciting proposition has come up in that our systems are being linked to Dr. Stanley Cohen's pharmacy based system on drug interaction because our languages are compatible. His system was also initially developed on ACNE equipment. Of course, he uses the hospital Business Office information in his pharmacy based system. We would use a shared data base with him as well as provide on-line quality control for the use of antibiotics. Hence, when drugs are ordered from the pharmacy prior to their issuance to the wards, the reports currently coming out of our Diagnostic Microbiology Laboratory would be used together with appropriate rules to advise all concerned as to their suitability.

It is quite likely that Dr. Howard Sussman's clinical chemistry information system will also be linked in the future to these systems to provide data on

potcutial limitations to use of antimicrobials which are an important part of the quality control of physician decision making. As you can see, having all three of these systems linked up to a common hospital base facility obviously allows interactive programs and shared data bases which would not be possible without much interfacting difficulties. Therefore, I believe a common hospital system will promote similar collaboration for others in the future.

Can you send me a copy of the application on Computer Health Care Applications Research for my files? Thank you.

$- 134 -$

APPENDIX B -

STANFORD UNIVERSITY HOSPITAL DATA PROCESSING DEPARTMENT

March 7, 1973

TO: Peter F. Carpenter, Assistant Vice President of Medical. :ffairs

FROM : $\sqrt{2}$, $\sqrt{2}$ **V. H. Barber, Assistant Controller for EDP** \subset \mathbb{R}^{N+1} is $\lambda_{\rm e} \sim \lambda_{\rm e}$

SUBJECT: Medical Center Computer Planning Chronology

Presented below is a chronology of events related to computer planning from late 1970 to date.

Late 1970 - Early 1971

Medical Center Sub-Committee for Computing accomplished very little except for a survey of computer and data processing needs at the Stanford University Medical Center.

October 1971

President's Computer Science Advisory Committee annual visit results in general observation that computer planning has dsteriorated,

December 8, 1971

Medical Center computer briefing to Dean Clayton Rich. Presentations by:

- V. Barber
- C. Dickens
- G. Frank1 in
- R. Jamtgaard
- T. Phillips
- M. Roberts

December 28, 1971

Medical Center Computer Planning Committee created.

Chai rman: E. Levinthal Members: S. Cohen, M.D. J. DeGrazia, M.D. E. Dong, M.D. S. Kalman, M.D. R. Jamtgaard T. Rindflcisch J. Stead J, Williams V. Barber

P. F. Carpenter March 7, 1973

Medical Center Computer Planning Committee meetings were held on:

1/24/72 Various configurations of computers, utilization of HDP and ACME l/31/72 loads were monitored. Organizational structures were studied; 2/15/72 Iong- and short-term plans were considered. Needs of research $3/6/72$ groups were put forth. First major report of hardware alterna-3/20/72 tives was presented March 20, 1972. k/10/72 4/24/7z 5/3/72 S/19/72 5/31/72 Various meetings in June

July 18 - August 3, 1972

Presentations of the various alternatives to computing in the Medical Center were made to the Computer Planning Comnittee.

- July 18 Position paper advocating PDP-10 for the Stanford Nedical Center Service Computing Facility - R. Jemtgaard and T. Rindfleisch.
- July 26 Stanford University Medical Center Proposed Service Facility position paper - V. Barber.
- August 3 Position paper advocating that computing service for the Stanford University Medical Center be supplied by a University computing facility - G. Franklin, T. Phillips, M. Roberts,

August 17, 12, 1972

Recap of Committee activity and alternatives for computing to Dean Rich. Made recommendation to him for computing. The conclusions of the Committee are attached in letters from E. Levinthal dated August 17 and August 18, 1972.

August 22, 1972

Medical School Executive Committee meeting. Clayton Rich, M.D., updated Executive Committee on computing alternatives (see attached letter of August 22, 1972).

August 21-23, 1972

 ${\tt Clayton}$ Rich dismissed original committee (see 12/28/71) and created an interim comni ttee:

P. F. Carpenter

Purpose: Summarize the financial and technical findings of the Medical Center Computer Planning Committee.

August 30, 1972

Interim Committee made its summarizing report to Clayton Rich (Copy attached: Letter of August 30, 1972).

September, 1972

Gene Franklin made recommendation to Vice Presidential Group regarding University-wide solution to computing. He was directed to draft a policy statement and a plan.

November 8, 1972

Arc Advisory Group on Computing Merger was established consisting of:

This group appointed a Planning Task Force made up of:

Chairman: C. Dickens
Members: V. Barber Members: R. Jamtgaard M. Ray F. Riddle

November - December 1972

Task Force has several sub-committees. Various meetings were hel during this period of time.

December 29, 1972

Task Force submitted its report and recommendation to the Advisory Group. Recommendations are attached.

January 1973

Dean Clayton Rich asked if the original SUH Data Processing proposal (see July 26, 1972) could offer a possible solution to Medical Center computing.

P. F. Carpenter March 7, 1973

January - February 1973

Numerous meetings and analyses were conducted in this period. Resul ts were a 360/65 or 370/158 if properly organized and planned could solve Medical Center computing needs.

February 23, 1973

Recommendation to Vice Presidential Group for purchase of a 370/158.

March 1973

Medicat Center computing solution still under study.

vhb: adg

APPENDIX C

(Excerpt from ACME Note HAD)

IBM 2701/SATELLITE COMPUTER MULTIPLEXOR DESIGN AND OPERATION

I. INTRODUCTION

This paper is intended to describe the design criteria, specifications and feature, theory of operation, and operational procedures for the IBM 2701/SATELLITE COMPUTER MULTIPLEXOR. The design criteria section explains some design philosophies and some desirable features that such a system should have. Features section gives a list of specifications and features. Theory of operation explains in detail how this system works. And finally operational procedures section gives detailed trouble shooting procedures for problem isolation and procedures to bring on a new user.

II. DESIGN CRITERIA

The purpose of a HOST/SATELLITE COMPUTER MULTIPLEXOR is to allow several remote satellite computers to communicate directly to a host computer and vice versa. The main function of the satellite computer multiplexor is to allow only one satellite computer to communicate to/from the host computer at a time. The satellite computer multiplexor should be capable of handling up to sixteen remote satellite computers. The satellite computer multiplexor should be designed such that it will be independent of the host computers' and satellite computer's designs and/or operational characteristics. All remote satellite computers, 100 feet away from the host computer, must transmit data serially to/from the host computer via the satellite computer multiplexor. The satellite computer multiplexor must be capable of timing out in the event of any malfunction or due to one particular satellite computer which has used up its allotted time in transmitting data to/from the host computer. And lastly, the host computer must be capable of interrupting any of the satellite computers via the satellite computer multiplexor.

In order to meet the above criteria, the satellite computer multiplexor can be thought of as made up of three basic sections: host computer interface, multiplexor control, and satellite computer I/O control, as shown in Figure 1.

The function performed by the host computer interface is handling all I/O signals to/from the host computer.

The functions performed by the multiplexor control are queueing satellite computer interrupt requests, establishing communication with the host computer, making sure that proper identification from the satellite computer is passed to the host computer, passing status to the host and to the satellite computer at all phases of the data transfer, detecting timeout conditions, monitoring and flaging any malfunction for troubleshooting purposes, and allowing the host computer to interrupt any satellite computer.

 \sim

FIGURE 1 SATELLITE COMPUTER MULTIPLEXOR

 $\sim 10^{-1}$

The functions performed by the satellite computer I/O control are serializing and deserializing data to/from remote satellite computers and allowing parallel data transfer if satellite computers are within 100 feet of the host computer. Serial data are to be transmitted bit asynchronous and an optional choice between word or character asynchronous or synchronous.

In order to maintain complete flexibility at the satellite computer end because of different computers, the interface between the satellite computer multiplexor and the satellite computer is to be divided into general and special interfaces. The general interface is to handle all I/O signals to/from the satellite computer multiplexor and the special interface is to handle all I/O signals to/from a particular satellite computer.

For implementation, the host computer is an IBM 360/50 and the satellite computer multiplexor is interfaced to one of the ports of an IBM 2701 Parallel Data Adapter (PDA). This means that the satellite computer multiplexor will work with any IBM 360/370 host computer as long as it is interfaced through an IBM 2701 PDA port. Remote satellite computers, on the other hand, can be DEC PDP-8, 9 , 10, 11, 12, 15 or XDS Sigma 3, 5, 7, or Hewlett/Packard HP-2411, 2115, 2116, or VArian 62Oi, 620f, or etc.

III. SPECIFICATIONS AND FEATURES

- 1. Handle up to 16 simultaneous satellite computers.
- 2. The satellite computer multiplexor is interrupt driven. It operates strictly on demand/response basis.
- 3. Each satellite computer talks to the IBM 360 on a first come, first served basis.
- 4. Each satellite computer can be assigned to any one of the 16 multiplexor channels.
- 5. Each satellite computer has a hardware key address at the satellite computer multiplexor end for ID purposes.
- 6. Transmission mode is by serial asynchronous half duplex for remote and/or parallel asynchronous for local operation.
- 7. Transmission speed is hardwired and the available speeds are: 250K*, lOOK*, 50K, lOK, 5K baud per second.
- 8. Word transmission rate for maximum word length (20 bits) is : 12.5K, 5K, 2.5K, 500, 250 words per second.

*Recommended for twisted pair less than 1000 feet or coaxial cable for longer distances
- 9. Maximum serial bit transmission between satellite computer multiplexor and satellite computer is 20 bits; that is 1 start bit, 2 control bits, 16 data bits, and 1 stop bit.
- 10. Maximum word length from satellite computer is 16 bits.
- 11. Data path between IBM 2701 and satellite computer multiplex is 16 bits wide.
- 12. The satellite computer has the option to run in complete demand/ response (synchronous by character) or semi-complete demand/ response (asynchronous by character) modes. Note this is not a programmable function.
- 13. The satellite computer running under complete demand/response mode requires four twisted pairs and operates at lower data rate.
- 14. The satellite computer running under semi-complete demand/ response mode requires only two twisted pairs and operates at higher data rate.
- 15. The IBM 360 asynchronously can interrupt any satellite computer via the multiplexor.
- 16. The IBM 360 can pass status to a satellite during the normal transmission cycle.
- 17. The satellite computer will receive all error and termination conditions through coded messages from the multiplexor so that it can act accordingly.
- 18. Detailed handshaking procedures between the satellite computer and the host computer are described in the section "Asynchronous/ Synchronous Data Transfer between Satellite and Host Computers".

ACME Note

ACNE Notes Index

M-40 Erica Baxter Mey 11, 1973

ACME Notes, written by all members of the ACME staff, are informal working papers. rney are divided below into four main categories: General Information, Administration
and Utilization, System Information, and User Information. Subcategories under System
Information and User Information parallel each oth and the ACME gtetisticsl Library are listed et the end of the Index.

The letters in the ACME Note codes are for filing and reference purposes only; the numbers in the codes--except for part of the J series--indicate reissue
All but historians can dispose of superseded issues. The J series and part of other ACME Notes are incorporated into the PL/ACME Manual (AM) revisions.

If you wish to have a copy of an ACME Note, it is available at the ACME office. Thoee notes preceded by en asterisk (*) are new or have been changed in some way since the last ACME Notes Index was issued.

ACNE Notes which have become OBSOLETE with this issue of AA are listed separately in the laet section to this index.

GENERAL INFORMATION

 $AA=40$

ADMINISTRATION AND UTILIZATION

SYSTEM INFORMATION

GENERAL

COMPILER AND LANGUAGE

WP1-4 ACME Subroutine PICK (Wiederhold/Berman)

 $\bar{\gamma}$

EXECUTOR

 $\sim 10^6$

INPUT/OUTPUT

TERMINALS AND DISPLAYS

FILES

AA-40

REAL-TINE

HARDWARE

 $AA-4O$

 $\mathbf{A}\mathbf{A}+\mathbf{\tilde{a}}_{\mathrm{V}}^{\mathrm{N}}$

OPERATING SYSTEM

 $-151 -$ 0S/360 Utility Program (Class/Miller) $QH = 1$ JUL 11, 1967 Job RXREF - Reverse Cross-Reference Program (Liere) $RX-1$ APR 22, 1970 TRMTYP-1 Format of ACME Control Block TRMTYP (Stainton) OCT 2. 1972 $117 - 7$ Instructions for Operating the 1800 Using the Time-Sharing System (Crouse) JUL 14, 1969 Notes to 1800 Operators (Crouse) $UOP-1$ $0.07 - 2 = 196.8$ $WFR-3$ FORTRAN-H Built-In Functions (Sanders) HIL 26. 1968 WNITE-1 Internal Documentation of Overnight TAPE and LISP Jobs (Granieri) FFB 13, 1973 $USC - 1$ Use of Stand-Alone Loader (Sanders) JAN 10, 1967 $WSCB-1$ 0S System Control Blocks (Frey)
SEP 12. 1972 ***usuc-3** ACME Written SVCs (Frey/Stainton) APR 20. 1973 YAA-3 Operations Procedures Table of Contents (Class) DEC 16, 1969 $YAS-1$ ACME Switchboard (Matous) JAN 26, 1971 $YRP-7$ OS/360 (MVT) Batch Processing Procedure (Class/Sutter) DEC 19 1972 YRTH-1 ACME IBM Equipment's Physical Data (Class) APR 11, 1969 $YCC-3$ 1052 Console Commands (Class/Sutter) DEC 19, 1972 $YCD-2$ Clip Deck--Change a Volume Label (Class) JAN 4. 1973 YCLEAN-2 Scratching Temporary Data Sets (Class) DEC 19, 1972 $YCOL-1$ Computer Operator's Lunch, Dinner Schedule (Matous) JAN 26, 1971 $YCON-I$ Converting Upper-Case Alpha Characters to Lower-Case with the Addition of a Punch (Class) APR 4, 1969 DASDI Disk Pack (Class) YDAS-3 DEC 20, 1972 $YDB-1$ Examples of Broadcasts and Debug Massages (Class) JUN 4, 1968 MVTDEBE - OS/360 Multi-Purpose Utility (Frey) YDEBE-1 DEC 7, 1971 DUMP/RESTORE ACME Operating System (Class) $YDR-2$ MAY 29, 1969 YEREP-2 Hardware Debugging Program--EREP (Class) APR 12, 1969 YHELP-1 What to Do in Case of an Irrecoverable Disk Error (Class) APR 21, 1969 YIBM-2 ACME's IBM Hardware (Class) DEC 28, 1970 YIBMTP-1 Operator Procedure: User Disk/Tape Dump/Restore Procedure (Germano) JAN 4, 1973 System IPL Procedures for Release 20.6 of OS/360 (MVT) (Class/Sutter) Y1PL-4 DEC 19, 1972 YLCS-1 Procedure to Test 2361 Bulk Core--LCS (Class) JUN 4, 1968 YLISP-1 Operator Procedures- Overnight LISP Jobs (Granieri/Class) FEB 6, 1973 $YLOG-1$ ACME 360 Console Log (Class) JUN 3, 1970 $YLPB-1$ Loading Printer Buffer (Class) JUN 4, 1968
Label Tapes (Class) $Y1.T-2$ DEC 19, 1972 $M-9$ ACME Monthly Accounting Programs ACCTTING and NEWACCT (S. Miller) OCT 12, 1972 $YMC-8$ Billing Cards for SCC (Miller) OCT 12, 1972 ACME Monthly Accounting Program: SUMMARY (Miller/Liere) $YMS - 5$ AUG 16, 1972 Loading NAMEs, PROJECTs, and DATA SETs on ACME (Class/Matous) $YNP-2$ APR 9, 1969 **YOAS-2** Operating the ACME System (Class/Sutter) OCT 21, 1970

YOUT-1 1800-360 Interface Recovery Instructions (Class) APR 11, 1968

- $YOX-1$ Procedure for 270X Error Restart (Denke)
- MAR 6, 1968
- $YPDF-1$ Loading the PDP-11 (Lew) NOV 16, 1972
- YPOW-1 Racovering From a Power Failure and Powering Off (Class) JUL 3, 1969

 $AA-4O$

USER INFORMATION

 $\mathcal{A}_{\mathcal{A}}$:

 $\mathcal{L}^{\text{max}}_{\text{max}}$ and $\mathcal{L}^{\text{max}}_{\text{max}}$

- 336 User Note: Sew Bulk Core and Disks, PLCTPRIN. Loma Linda Graphics. New Commands in PL/ACME (Staff) NOV 5, 1971
- 537 User Note: Perfomnce Tests on New Bulk Core. New Text Editing Features. DATACOPY, EVENT, User Tape Services, New Character Scriag Function, New Information on LOGOFF (Staff) NOV 18. 1971
- 538 User note: Holiday Schedule, Permitting of Real Time Lines, User Services, Consulting Service Schedule, Clean Up Your Files! (Staff) DEC 21. 1971
- 539 User note: Seminar on Time-Oriented Medical Records, Small Machine Multiplexor, Medical Center Computer Facility Planning Committee, Printing and Punching Services, Double Precision Argument Bug, New Show Command: SHOW DSOPEN (Staff) FEB 7, 1972
- 540 User note: Revised Operating Hours, Medical Computing Seminars, PL/1 versus FL/ACME Compatibility, Small Machine Multiplexor Interface, System Errors 226 and 237(File System), Program CATALOG PUBLIC, Fast Fourier Transforms, Antilogs (Staff) FEB 28, 1972
- 341 User note: Revised Rates for ACME Service, New Terminal Support (Memorex), Medical Computing Seminars, Reassignment of Gio Wiederhold, New ACME Note Index, New LISTAKER/JOBTAKER Feature, Bug: Character String Variable and Title Option, Card and Paper Recycling, Acknowledging ACMB In Publications (Staff)
	- APR 10, 1972
- $.142$ User note: Medical Center Computer Facilities Planning Committee, Small Machine Multiplexor Schedule, 30 Char/Sec Typewriter Terminals, 120 Char/Sec Alphanumeric CRT Service, SUMEX Research Proposal, Comments on Lack of Core Storage Error Messages. New Version of PUBLIC Program COPIER. User Listings Schedule, PDP-11 Simulator, New Versions of PUBLIC Programe LINREG and LACKFIT (Staff) JUN 13, 1972
- 543 User note: Application of the ACME Quantity Discount to More Than One ACME Account, New PUBLIC Program RECOMPOS: Recomposing PL/ACME Programs from Card Decks, Error in ACMB Note ELVSIM: PDP-11 Simulator on ACME, ACME Staff Directory (Staff) JUL 11, 1972
- 544 User note: Medical Center Computer Facilities Planning Committee. 30 Char/Sec Typewriter Terminal Selection, Transfer of Lee Hundley to SLAC, ACME Grant Status, New ACME Statement and Command to Adjust Width of Output (Staff) AUG 9. 1972
- **J45** User note: 30-Character-Per-Second Terminal Availability, Frank Germane-New User Services Manager, SHOW FREQUENCY-Debugging/Optimization Aid, New PUBLIC Programs BIBLIO and BIBUP: Interactive Keyword/Entry Programs, New PUBLIC Program CONCORD: Generates Concordance Files from Text Files, New PUBLIC Program ARSPEC: Spectral Density Analysis, New PUBLIC Program SURVIVR: Life Table Construction, LISP Bug, Error in PUBLIC Program TSQUARE. Error in ACME Note FDFORM-1 (Format of ACME Dump Tepcs for Users) (Staff) SEP 14. 1972
- **J46** User note: Do You Have Data Storage and Retrieval Problems?. Pilot Study/Implementation of Public Data Bank Programs, Time Oriented Data Bank Seminar, Comment on Pageminute ACCOUnting. G.E. Terminet Terminals, Terminal Demonstration, New PUBLIC Programs: QSORT, QBANR, and SORTEXT - Sorting Routines, New PUBLIC Program TEACHER: Teaches PL/ACMB, "Super ACME Index. Recovery from Realtime Input/Output Failures (Staff) OCT 23, 1972
- 547 User note: Medical Center Computing: What is going to happen? (Staff) Announcement of InfOrmational Meeting NOV 10, 1972
- 548 User note: Index to the ACME PUBLIC Programs. New PUBLIC Program IBMTAPE: Enters Requests to Dump/Restore Users' ACME Files to/from Tape, ACME Holiday Schedule, Radioimmunoassay Users, New PUBLIC Program ARITHDAT: Arithmetic Date Routines, New PUBLIC Programs BSORT and BRANK, SORTEXT Users, Correspondence Terminal Support Changes (Staff) DEC 20. 1972
- 349 User note: New Medical Center Computing Facility, ACME Datafile Compression, Revised ACME Operator Schedule, Comments from ACME Users, SIGBIO Colloquium on Time-Oriented Medical Records, New Seminar Series. Personnel Changes, Changes in PUBLIC Program DATACOPY, Use of LF or INDEX Key, New and Revised Linear Regression Programs on PUBLIC File, Long Print Job Problem (Staff) FEB 13, 1973
- *550 User note: Medical Center Computer Facility, TOD Seminar, PL/ACME Classes, LISP Available (Staff) MAR 8, 1973
- KCC-1 PL/ACME Collating Sequence (Wiederhold/Emmona) MAY 12, 1969
- NI-2 Initialization of Variables (Hundley)
- SEP 27, 1969
- NSC- 1 Subscripting Cost (C. Wiederhold) AUG 17, 1972

 $-155 -$

- JAN 31, 1972
String Handling Functions (Wiederhold) $TS-7$
- $TST-1$
- String Handling Functions (Wiederhold)
SEP 12, 1969
JAN 13, 1970
JAN 13, 1970
Options for GET and PUT Statements to Allow In Memory Conversion
(Wiederhold)
JAN 7, 1971 $TW-1$
-
- Proposal for Arithmetic on Strings (Wiederhold)
SEP 4, 1970 $WDB-1$

INPUT/OUTPUT

TERMINALS AND DISPLAYS

-
-
- OCT 16, 1972
TT-1 Special Character String Input/Output Procedures (Sanders)
FEB 1, 1968
WBEE-2 Using the Beehive Terminal On ACME Through the PDP-11 (Briggs)
NOV 4, 1971
WPDPB-1 Using the Litton Printer On ACME Through th

FILES

 $\mathcal{L}_{\mathcal{L}}$

 $AA-40$

 $\label{eq:2.1} \frac{1}{2} \sum_{i=1}^n \frac{1}{2} \sum_{j=1}^n \frac{$

- 158 -

APR 4. 1973

RFAL-TIME

 \sim

 $\Lambda\Lambda\!\!=\!\mathbf{\mathbf{4.3}}$

ACME SUBROUTINE LIBRARY

 $AA-40$

- MAR 12. 19/l
- FAV-4 ACME Subroutine MATPRD: Multiplies Two Matrices (Liere) SEP 5, 1969
- $RAW-3$ ACME Subroutine SMOOTH: Smooths Time Series (Liere/Whitner)
- EAX-3 FER 4, 1970 ACHE Subroutine CLRATIO: Confidence Limite for Ratio of Two Means (Liere)
- SEP 12, 1969
- FAY-1 ACME SubroutineORDER: Dependent and Independent Variables from Syrmaetric Correlation Matrix (Liere) SEP 12, 1969
- EAZ-3 ACME Subroutine TRANSPOSE (Whitner) APR 8. 1970
- $EBF-3$ ACME Subroutine MATADD: Adds Two Matrices (Liere) .IAN 24. 1969
- t:BT-2 ACME Subroutine EIGEN: Calculates Eigenvalues and Eigenvectors for a Symmetric Matrix (Liere) MAY 26. 1969
- $EBZ-2$ ACME SubroutineRUNGA: First-Order Differential Equations by Runga-Kutte Kethod (G. Sanders) FEB 17, 1969
- ED-3 Statistical Library Testing Programs (Liere/Whitner) AFR 7, 1970
- EDH-2 ACME Subroutine NULTR: Hultiple Linear Regression (Liere/Whitner) .JUL 19. 1971
- $EDO-2$ ACXI? Subroutine WALT: Noves Half-Word Integer Data from One Array to Another (Rreitbard) AUG 11, 1969
- $EDW-1$ ACME SubroutIne PLOTLINE: Plots Curve on Designated Digital Plotting Device (G. Sanders) OCT 9. 1969
- EDX-1 ACME Subroutine PLOTTS: Plots Time Series on Designated Digital Plotting Device (G. Sanders) SEP 12, 1969
- EEC- 1 ACME Subroutine WSRRG: Simple Linear Regression for Weighted or Unweighted Data Whitner) DEC 15, 1970
- EEW-1 Accuracy of ACME Statistical Subroutines (Whitner) AUC 9. 1971
- EEX-1 Usage of ACME Statistical Subroutines (Whitner) JUL 30. 1970
- $EZY-1$ ACME Subroutine CODE: Encodes and Decodes User's Piles (Hale/Whitner/Nye) DEC 16. 1970

ACME PROGRAM LIBRARY - MOST ON PUBLIC FILE

 $\Delta A = h(r)$

IUN 25, L971

APPENDIX E USER PUBLICATIONS

1. Published in Regularly Scheduled-periodicals During Fiscal 1973

- Alderman, E.L., W.H. Barry, A.F. Graham, and D.C. Harrison, an, E.L., W.H. Barry, A.F. Graham, and D.C. Harrison, "Hemodynamic Effects of Morphine and Pentazocin
Differ in Cardiac Patients," NEW ENGLAND JOURNAL OF MEDICINE, Vol. 287, pp. 623-627, Sept. 28, 1972, NEW ENGLAND JOURNAL OF MEDICINE, Vol. 287, pp. 623-627, Sept. 28, 1972.
- Alderman, E.L., A. Branzi, S. Sanders, D.C. Harrison, and B.W. Brown, "Evaluation of the Pulse Contour Method of Determining Stroke Volume in Man," CIRCULATION, Vol. XLVI, September 1972.
- Barry, W.H., A.M. Marlon, and D.C. Harrison, "The Hemodynamic Effects of Strontium Chloride in the Intact Dog," PROCEEDINGS OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY AND MEDICINE, Vol. 141, No. 1, pp. 52-58, October 1972.
- Branzi, A., H. Mailhot, E.L. Alderman, and D.C. Harrison, "Ultrasound Determination of Left Ventricul Position for Volume Angiography," CHEST, Vol. 62, pp. 29-33, July 1972.
- Brown, C.R., P.F. Shroff, and W.H. Forrest, Jr., "Relative Potency of Triclorofos Compared to Pentobarbit as a Hypnotic," J. OF CLIN. PHARM. AND NEW DRUGS, Vol. 12, No. 8 and 9, pp. 306-312, August-September 1972.
- Brown, C.R., W.H. Forrest, Jr., and H. Hayden, "The Respiratory Effects of Pentobarbital and Secobarbital in Clinical Doses," JOURNAL OF CLINICAL PHARMACOLOGY, Vol. 13, pp. 28-35, January 1973.
- Buchanan, B.G., E.A. Feigenbaum, and N.S. Sridharan, "Heuristic Theory Formation: Data Interpretation and Rule Formation," MACHINE INTELLIGENCE 7, Edinburgh University Press, pp. 267-290, November 1972.
- Forrest, W.H., Jr., P.F. Shroff, and D.L. Mahler, "Analgesic and Other Effects of Nalmexone in Man," J. OF CLIN. PHARM. AND THERAP., Vol. 13, No. 4, pp, 520-525, July-August 1972.
- Forrest, W.H., Jr., C.R. Brown, P.F. Shroff, and G. Teutsch, "Relative Potency of Propiram and Morphine for Analgesia in Man," JOURNAL OF CLINICAL PHARMACOLOGY, Vol. 12, No. 11 & 12, Nov.-Dec. 1972.
- Goldman, R.H., R.N. Deutscher, E. Schweizer, and D.C. Harrison, "Effect of a pharmacologic dose of digoxin on inotropy in hyper- and normokalemic dogs," AMERICAN JOURNAL OF PHYSIOLOGY, Vol. 223, No. 6, pp. 1438-1443, December 1972.
- Herman, M.M., K.G. Sensch, K.W. Marich, and D. Glick. "The Effects of Gold Thioglucose on Mouse Fibroblasts in vitro: Morphological and Laser Microprobe Studies", EXPERIMENTAL AND MOLECULAR PATHOLOGY, Vol. 16, No. 2, pp. 186-200, April 1972.
- Joffe, J.M., K. Milkovic, and S. Levine, "Effects of Changes in Maternal Pituitary-Adrenal Function on Behavior of Rat Offspring," PHYSIOL. BEHAV., Vol. 8, pp. 425-430, 1972.
- Lamb, E., and A.L. Cruz. "Data Collection and Analysis in an Infertility Practice," FERTILITY AND STERILITY, Vol. 23, No. 5, pp. 310-319, May 1972.
- Lamb, E., "Prognosis for the Infertile Couple," FERTILITY AND STERILITY, Vol. 23, No. 5, pp. 320-325, May 1972.
- McConnell, H. M., and P. Devaux, "Lateral Diffusion in Spin Labeled Phosphatidyl Choline Multilayers," J. AM. CHEM. sot., Vol. 94, p. 4475, 1972.
- Marich, K.W., J.B. Orenberg, W.J. Treytl, and D. Glick, "Health Hazards in the Use of the Laser Microprobe for Toxic and Infective Samples," AMERICAN INDUSTRIAL HYGIENE JOURNAL, pp. 488-491, July 1972.
- Northway, W.H., R. Petriceks, and L. Shahinian, "Quantitative Aspects of Oxygen Toxicity in the Newborn: Inhibition of Lung DNA Synthesis in the Mouse," PEDIATRICS, Vol. 50, No. 1, pp. 67-72, July 1972.
- Raybin, D.M., L.L. Bertsch, and A. Kornberg, "A Phospholipase in Bacillus megaterium Unique to Spores and Sporangia," BIOCHEMISTRY, Vol. 11, No. 10, pp. 1754-1760, 1972.
- Saffir, A.J., K.W. Marich, J.B. Orenberg, and W.J. Treytl, "Statistical Evaluation of Background and Laser Energy Corrections of Spectral Line Signal in Laser Microprobe Analysis", APPLIED SPECTROSCOPY, Vol. 26, No. 4, pp. 469-471, 1972.
- Saffir, A.J., and H.M. Myers, "A Modern Statistics Course for Dentists Using a Time-Sharing Computer", JOURNAL OF DENTAL EDUCATION, April 1972.
- Scandella, C.J., P. Devaux, And H.M. McConnell, "Rapid Lateral Diffusion of Phospholipids in Rabbit Sarcoplasmi Reticulum," PROC., NAT. ACAD. SCI., Vol. 69, No. 8, pp. 2056–60, 1972.
- Smith, D.H.. B.G. Buchanan, R.S. Engelmore, A.M. Duffield, A. Yeo, E.A. Feigenbaum, J. Lederberg, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference VIII. An Approach to the Computer Interpretation of the High Resolution Mass Spectra of Complex Molecules. Structure Elucidiation of Estrogenic Steroids," JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, Vol. 94, pp. 5962-5971, 1972.
- Treytl, W.J., Orenberg, J.B., Marich, K.W., Saffir, A.J., and Glick, D., "Detection Limits in Analysis of Metals in Biological Materials by Laser Microprobe Optical Emission Spectrometry", ANALYTICAL CHEMISTRY, Vol. 44, p. 1903, September 1972.
- Weichsel, M.E., Jr., N.J. Hoogenraad, R.L. Levine, and N. Kretchmer, "Pyrimidine Biosynthesis During Development of Rat Cerebellum," PEDIAT. RES., Vol. 6, pp. 682-686, 1972.

2. Other Publications

- Brown, B.W., Spurious Appearance of Mosaicism in Three Generations in One Family with a 3/B Translocation," with B. Efron and R.G. Miller, submitted for publication (spring 1973).
- .Brown, C.R., W.H. Forrest, Jr., and J. Hayden, "Respiratory Effects of Hydromorphone in Man," CLIN. PHARM. AND THERAP., accepted for publication (October 1972).
- Brown, C.R., P.F. Shroff, and W.H. Forrest, Jr., "The Oral Hypnotic Bioassay of Hydroxazine and Pentobarbit for Night-time Sedation," in preparation (winter 1972).
- Cohen, S.N., L. Crouse, M.F. Armstrong, and G.Hunn, "A Computer-Based System for Prospective Identification of Drug Interactions", PROCEEDINGS OF THE FIFTH INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, Univ. of Hawaii, 1972.
- Cohen, S.N., M.F. Armstrong, L. Crouse, and G. Hunn, "A Computer-Based System for Prospective Detection and Prevention of Drug Interactions", DIA BULLETIN, in press (fall 1972).
- Corby, James C., and B.S. Kopell, "Differential Contribution of Blinks and Vertical Eye Movements as Artifacts in EEG Recording," PSYCHOPHYIOSLOGY, in press (winter 1972).
- Corby, J.C., and B.S. Kopell, "The Effect of Predictability on Evoked Response Enhancement in Intramodal Selective Attention," PSYCHOPHYSIOLOGY, in press (winter 1972).
- Devaux, P., C.J. Scandella, and H.M. McConnell, "Spin-Spin Interactions Between Spin Labeled Phospholipids Incorporated into Membranes," submitted to the JGURNAL OF MAGNETIC RESONANCE (winter 1972).
- Devaux, P, C. Scandella, and H.M. McConnell, "Etude par Marquage de Spin de la Diffusion Laterale Dans Les Membranes Biologiques," in press (winter 1972).
- Dingledine, R., and A. Goldstein, "Lethality of the Opioid Narcotic Levorphanol," BRIT. J. PHARMACOL, in press (spring 1973).
- Dingledine, R., and A. Goldstein, "Lethality of the Morphinan Isomers Levorphanol and Dextrorphan," BRIT. J. PHARMACOL., in press (spring 1973).
- Forrest, W.H., Jr, C. R. Brown, R. Delfalque, J. Katz, D.L. Mahler, P. Shroff, and G. Teutsch, "Report of the Veterans Administration Cooperative Analgesic Study," BULLETIN - PROBLEMS OF DRUG DEPENDENCE, Vol. 1, October 1972.
- Forrest, W.H., Jr., C.R. Brown, and J. Belleville. "Studies of Premedicants with Emphasis on Methodology: Part I - Barbiturates," FIFTH WORLD CONGRESS OF ANAESTHESIOLOGISTS, abstracts of papers presented, September 19, 1972.
- Forrest, W.H., Jr., and C.R. Brown, "The Effects of Ethyl Alcohol, Delta-9-THC and Pentobarbital on a Tracking Task," FIFTH WORLD CONGRESS OF ANAESTHESIOLOGISTS, abstracts of papers presented, September 19. 1972.
- Forrest, W.H., J.W. Bellville, and B.W. Brown, "The Interaction of Caffeine with Pentobarbital as a Nightime Hypnotic," submitted for publication (spring 1973).
- Fries, J., "The Effect of Ice Water on Esophageal Rewarming in Connective Tissue Diseases (CTD)," abstract submitted to GASTROENTEROLOGY (winter 1972).
- Fries, J., and R. Siegal, "Testing the Preliminary Criteria for SLE," ANN. RHEUMATIC DISEASE, in press (winter 1972).
- Goldstein, A., and B.A. Judson, "Three Critical Issues in the Management of Methadone Programs," Chapter in ADDICTION: A COMPREHENSIVE TREATISE, editor P.G. Bourne, Academic Press, 1973.
- Goldstein, A., and B.A. Judson, "Efficacy and Side Effects of Three Widely Different Methadone Doses,' PROC. FIFTH NATIONAL CONF. ON METHADONE TREATMENT, in press (spring 1973).
- Leiderman, P.H., A.D. Leifer, M.J. Seashore, C.R. Barnett, and R. Grobstein, "Mother-Infant Interaction: Effects of Early Deprivation, Prior Experience and Sex of Infant," PROCEEDINGS OF THE 51ST ANNUAL MEETING OF THE ASSOCIATION FOR RESEARCH IN NERVOUS AND MENTAL DISEASE, Baltimore, Williams and Wilkins Co., 1972.
- Leifer, A.D., P.H. Leiderman, C.R. Barnett, and J.A. Williams, "Effects of Mother-Infant Separation on Maternal Attachment Behavior," CHILD DEVELOPMENT, in press (winter 1972).
- Knight, L., and L. Luzzatti, "The Replication Pattern of the X and Y Chromosomes in Partially Synchronized Lymphocyte Cultures," CHROMOSOMA, in press (winter 1972).
- McConnell, H.M., P. Devaux, and C. Scandella, "Lateral Diffusion and Phase Separations in Biological Membranes" BIOLOGICAL MEMBRANES, California Membrane Conference, pp. 27-37, 1972.
- Reynolds, W.C., "Mass Spectrometer Data Acquisition and Processing Systems, High and Low Resolution Mass Spectrometers, ' in BIOCHEMICAL APPLICATIONS OF MASS SPECTROMETRY, G.R. Waller (ed.), Wiley Interscience Press, p. 109, 1972.
- Rosenthal, W. S., and J. Eisenson, Longitudinal Studies of Linguistically Deviant Children," Final Report, Grant NS07514, National Institute of Neurological Diseases and Stroke, April 1972.
- Savage, I.R., and B.W. Brown, "Statistical Studies in Prediction of Attendance for a University," to be published in a volume on "Analytical Models for Educational Planning, ' edited by Hector Correa (spring 1973).
- Seashore, M.J., A.D. Leifer, C.R. Barnett, and P.H. Leiderman, "The Effects of Denial of Early Mother Infant Interaction on Maternal Self-Confidence," JOURNAL OF PERSONALTIY AND SOCIAL PSYCHOLOGY, in press (winter 1972).
- Teutsch, G., 0. Mahler, C.R. Brown, and W.H. Forrest, Jr., "A Study of the Hypnotic Efficacy of Dipenhydramine, Methapyrilene, and Pentobarbital, in preparation (winter 1972).
- Turnbull. B.W.. and B.W. Brown. "Survivorship Analysis of Heart Transplant Data," STANFORD UNIVERSITY TECHNICAL REPORT, NO. 34, October 4, 1972. -
- Wiederhold, G., "A Choice of Language to Support Medical Research," presented at the 1972 ACM Conference, Boston.
- Wiederhold, G., "Das ACME-System an der Stanford School of Medicine und seine praktische Verwenaung," KRANKENHAUS-INFORMATIONSSYSTEME, pp. 139-147, F.K. Schattauer Verlag, Stuttgart-New York, 1972.