S U M E X

STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE

RR - 00785

ANNUAL REPORT - YEAR 05

Submitted to

BIOTECHNOLOGY RESOURCES PROGRAM
NATIONAL INSTITUTES OF HEALTH

May, 1978

DEPARTMENT OF GENETICS
STANFORD UNIVERSITY SCHOOL OF MEDICINE

Joshua Lederberg, Principal Investigator
Edward A. Feigenbaum, Principal Investigator

Table of Contents

Section . . . Page

List of Figures iv

1. RESOURCE IDENTIFICATION . . . .

2. RESOURCE OPERATIONS 2

2.1 PROGRESS 2

2.1.1 RESOURCE SUMMARY AND GOALS 2

2.1.2 TECHNICAL PROGRESS 7

2.1.2.1 FACILITY HARDWARE 7

2.1.2.2 SYSTEM SOFTWARE 12

2.1.2.3 NETWORK COMMUNICATION FACILITIES 16

2.1.2.4 SYSTEM RELIABILITY AND BACKUP 22

2.1.2.5 SOFTWARE EXPORT 22

2.1.2.6 USER SOFTWARE AND INTRA-COMMUNITY COMMUNICATION . . 32

2.1.2.7 DOCUMENTATION AND EDUCATION 32

2.1.2.8 SOFTWARE COMPATIBILITY AND SHARING 33

2.1.3 RESOURCE MANAGEMENT 34

2.1.3.1 ORGANIZATION 34

2.1.3.2 MANAGEMENT COMMITTEES 34

2.1.3.3 NEW PROJECT RECRUITING 35

2.1.3.4 STANFORD COMMUNITY BUILDING 37

2.1.3.5 AIM WORKSHOP SUPPORT 37

2.1.3.6 RESOURCE ALLOCATION POLICIES 38

2.1.4 FUTURE PLANS 40

TABLE OF CONTENTS

2.2 SUMMARY OF RESOURCE USAGE 42

2.2.1 SYSTEM LOADING 42

2.2.2 SYSTEM EFFICIENCY 44

2.2.3 RELATIVE SYSTEM LOADING BY COMMUNITY 46

2.2.4 INDIVIDUAL PROJECT AND COMMUNITY USAGE 50

2.2.5 NETWORK USAGE 55

2.3 RESOURCE EQUIPMENT SUMMARY 58

2.4 PUBLICATIONS 59

3. RESOURCE FINANCES 60

3.1 BUDGETARY MATERIALS 60

3.2 RESOURCE FUNDING 60

4. COLLABORATIVE PROJECT REPORTS 61

4.1 NATIONAL AIM PROJECTS 62

4.1.1 ACQUISITION OF COGNITIVE PROCEDURES 63

4.1.2 CHEMICAL SYNTHESIS PROJECT 69

4.1.3 HIGHER MENTAL FUNCTIONS PROJECT 79

4.1.4 INTERNIST PROJECT 82

4.1.5 MEDICAL INFORMATION SYSTEMS LABORATORY 86

4.1.6 PUFF/VM PROJECT 93

4.1.7 RUTGERS COMPUTERS IN BIOMEDICINE 102

4.1.8 SIMULATION OF COMPREHENSION PROCESSES 117

4.2 STANFORD PROJECTS 122

4.2.1 AIHANDBOOK PROJECT 123

4.2.2 DENDRAL PROJECT 126

4.2.3 AGE PROJECT 143

4.2.4 HYDROID PROJECT 150

TABLE OF CONTENTS

4.2.5 MOLGEN PROJECT 156

4.2.6 MYCIN PROJECT 163

4.2.7 PROTEIN STRUCTURE PROJECT 177

4.3 PILOT AIM PROJECTS 182

4.3.1 COMMUNICATION ENHANCEMENT PROJECT 183

4.3.2 COMPUTERIZED PSYCHOPHARMACOLOGY ADVISOR 190

4.3.3 ORGAN CULTURE PROJECT 200

4.4 PILOT STANFORD PROJECTS 206

4.4.1 GENETICS APPLICATIONS PROJECT 207

4.4.2 QUANTUM CHEMICAL INVESTIGATIONS 212

Appendix I

AIHANDBOOK OUTLINE 217

Appendix II

DIGITAL COMMUNICATIONS AND THE CONDUCT OF SCIENCE -
THE NEW LITERACY 223

Appendix III

COMPARISON OF MAINSAIL AND PASCAL 235

Appendix IV

AIM MANAGEMENT COMMITTEE MEMBERSHIP 245

TABLE OF CONTENTS

List of Figures

1. SUMEX-AIM Computer Configuration 9

2. TYMNET Network Map 18

3. ARPANET Geographical Network Map 19

4. ARPANET Logical Network Map 20

5. TELENET Network Map 21

6. CPU Time Consumed by Month 42

7. Peak Number of Jobs by Month 43

8. Peak Load Average by Month 43

9. System Overhead by Month 44

10. Page Trap Time by Month 45

11. Monthly CPU Usage by Community 47

12. Monthly File Space Usage by Community 48

13. Monthly Terminal Connect Time by Community 49

14. TYMNET Usage Data 55

15. ARPANET Usage Data 56

NATIONAL INSTITUTES OF HEALTH
DIVISION OF RESEARCH RESOURCES
BIOTECHNOLOGY RESOURCES PROGRAM

SECTION I - RESOURCE IDENTIFICATION

Report Period: From August 1, 1977 to July 31, 1978		Grant Number: RR-00785 Report Prepared: May, 1978
Name of Resource: Stanford University Medical EXperimental Computer (SUMEX)	Resource Address: Stanford University Stanford, California 94305	Resource Telephone Number: (415) 497-5141
Principal Investigators: Joshua Lederberg, Ph.D. Edward A. Feigenbaum, Ph.D.	Titles: Chairman and Professor Chairman and Professor	Academic Departments: Department of Genetics School of Medicine Department of Computer Science
Grantee Institution: Stanford University	Type of Institution: Private University	Investigators' Telephone Nos.: Lederberg: (415) 497-5801 Feigenbaum: (415) 497-4079

Name of Institution's Biotechnology Resource Advisory Committee:
SUMEX-AIM Executive Committee

Membership of Biotechnology Resource Advisory Committee:

NAME	TITLE	DEPARTMENT	INSTITUTION
Saul Amarel, Ph.D.	Chairman & Professor	Computer Science	Rutgers University
Stanley Cohen, M.D.	Head	Div. Clinical Pharmacology	Stanford University School of Medicine
	Professor	Medicine,Genetics	
Donald Lindberg, M.D.	Professor	Pathology	University of Missouri School of Medicine
	Director	Information Science Group	University of Missouri School of Medicine
	Director	Health Care Technology Ctr.	University of Missouri - Columbia
Jack Myers, M.D.	University Professor of Medicine	At Large	University of Pittsburgh School of Medicine

Principal Investigators:	Signature:	Date:
Joshua Lederberg, Ph.D. Chairman and Professor	Joshua Lederberg	May 25, 1978
Edward A. Feigenbaum, Ph.D. Chairman and Professor	Edward A Feigenbaum	May 23, 1978
Stanford University Official: Janet P. Johnson Sponsored Projects Officer	Janet P. Johnson	5/24/78

SUMEX-AIM Resource Progress Report - Year 05

This annual report covers work performed under NIH Biotechnology Resources Program grant RR-785 supporting the Stanford University Medical EXperimental computer (SUMEX) research resource for applications of Artificial Intelligence in Medicine (AIM). It spans the year from May 1977 - April 1978.

2 RESOURCE OPERATIONS

2.1 PROGRESS

2.1.1 RESOURCE SUMMARY AND GOALS

The SUMEX-AIM project is a national computer resource with a dual mission: a) the promotion of applications of artificial intelligence (AI) computer science research to biological and medical problems and b) the demonstration of computer resource sharing within a national community of health research projects. The SUMEX-AIM resource is located physically and administratively in the Stanford University Medical School and serves as a nucleus for a community of medical AI projects at universities around the country. SUMEX provides computing facilities tuned to the needs of AI research and communication tools to facilitate remote access, inter- and intra-group contacts, and the demonstration of developing computer programs to biomedical research collaborators.

Artificial Intelligence research is that part of Computer Science concerned with the symbol manipulation processes that produce intelligent action (1). By "intelligent action" is meant an act or decision that is goal-oriented, is arrived at by an understandable chain of symbolic analysis and reasoning steps, and utilizes knowledge of the world to inform and guide the reasoning.

Some scientists view the performance of complex symbolic reasoning acts by computer programs as the sine qua non for artificial intelligence programs, but this is necessarily a limited view.

---

(1) For recent reviews to give some perspective on the current state of AI, see: (i) Boden, M., "Artificial Intelligence and Natural Man," Basic Books, New York, 1977; (ii) Feigenbaum, E.A., "The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering," Proceedings of the Fifth International Conference on Artificial Intelligence, 1977; (iii) Winston, P.H., "Artificial Intelligence", Addison-Wesley Publishing Co., 1977; and (iv) Nilsson, N.J., "Artificial Intelligence", Information Processing 74, North-Holland Pub. Co. (1975). An additional overview of research areas and techniques in AI is being developed as an "Artificial Intelligence Handbook" under Professor E. A. Feigenbaum by computer science students at Stanford (see page 123 for a status report and Appendix I for a current outline).

Another view unifies AI research with the rest of computer science. It is a simplification, but worthy of consideration. The potential uses of computers by people to accomplish tasks can be "one-dimensionalized" into a spectrum representing the nature of the instructions that must be given the computer to do its job; call it the WHAT-TO-HOW spectrum. At the HOW extreme of the spectrum, the user supplies his intelligence to instruct the machine precisely HOW to do his job, step-by-step. Progress in computer science may be seen as steps away from that extreme "HOW" point on the spectrum: the familiar panoply of assembly languages, subroutine libraries, compilers, extensible languages, etc. illustrate this trend.

At the other extreme of the spectrum, the user describes WHAT he wishes the computer, as his instrument, to do for him to solve a problem. He wants to communicate WHAT is to be done without having to lay out in detail all necessary subgoals for adequate performance yet with a reasonable assurance that he is addressing an intelligent agent that is using knowledge of his world to understand his intent, complain or fill in his vagueness, make specific his abstractions, correct his errors, discover appropriate subgoals, and ultimately translate WHAT he wants done into detailed processing steps that define HOW it shall be done by a real computer. The user wants to provide this specification of WHAT to do in a language that is comfortable to him and the problem domain (perhaps English) and via communication modes that are convenient for him (including perhaps speech or pictures).

The research activity aimed at creating computer programs that act as "intelligent agents" near the WHAT end of the WHAT-TO-HOW spectrum can be viewed as the long-range goal of AI research. Historically, AI research has been the primary vehicle for progress toward this objective, although a substantial part of the applied side of computer R&D has related goals, if an often fragmented approach. Unfortunately, workers in other scientific disciplines are generally unaware of the role, the goals, and the progress in AI research. Currently authorized projects in the SUMEX community are concerned in some way with the design of "intelligent agents" applied to biomedical research. The tangible objective of this approach is the development of computer programs which, using formal and informal knowledge bases together with mechanized hypothesis formation and problem solving procedures, will be more general and effective consultative tools for the clinician and medical scientist. The systematic search potential of computerized hypothesis formation and knowledge base utilization, constrained where appropriate by heuristic rules, empirical data, or interactions with the user, has already produced promising results in areas such as chemical structure elucidation and synthesis, diagnostic consultation, and mental function modeling. Needless to say, much is yet to be learned in the process of fashioning a coherent scientific discipline out of the assemblage of personal intuitions, mathematical procedures, and emerging theoretical structure of the "analysis of analysis" and of problem solving. State-of-the-art programs are far more narrowly specialized and inflexible than the corresponding aspects of human intelligence they emulate; however, in special domains they may be of comparable or greater power, e.g., in the solution of formal problems in organic chemistry or in the integral calculus.

An equally important function of the SUMEX-AIM resource is an exploration of the use of computer communications as a means for interactions and sharing between geographically remote research groups engaged in biomedical computer science research. This facet of scientific interaction is becoming increasingly important with the explosion of complex information sources and the regional specialization of groups and facilities that might be shared by remote researchers (see Appendix II on page 223). Our community building role is based upon the current state of computer communications technology. While far from perfected, these new capabilities offer highly desirable latitude for collaborative linkages, both within a given research project and among them. Several of the active projects on SUMEX are based upon the collaboration of computer and medical scientists at geographically separate institutions; separate both from each other and from the computer resource. The network experiment also enables diverse projects to interact more directly and to facilitate selective demonstrations of available programs to physicians, scientists, and students. Even in their current developing state, communication facilities enable effective access to the rather specialized SUMEX computing environment and programs from a great many areas of the United States (even to a limited extent from Europe). In a similar way, the network connections have made possible close collaborations in the development and maintenance of system software with other facilities.

As we complete the first 5-year term of the SUMEX-AIM resource grant, we can report that our initial technical task has been achieved. We have collected and implemented an effective set of hardware and software tools to support the development of large, complex AI programs and to facilitate communications and interactions between user groups. We have substantially increased the roster of user projects (from an initial 5) to 15 current major projects plus a group of pilot efforts. Many of these projects are built around the communications network facilities we have assembled; bringing together medical and computer science collaborators from remote institutions and making their research programs available to still other remote users. As discussed in the sections describing the individual projects, a number of the computer programs under development by these groups are maturing into tools increasingly useful to the respective research communities. The demand for production-level use of these programs has surpassed the capacity of the present SUMEX facility and has raised the general issues of how such software systems can be optimized for production environments, exported, and maintained.

A number of significant events and accomplishments affecting the SUMEX-AIM resource occurred during the past year:

1) Professor Lederberg has been the principal investigator and chairman of the SUMEX-AIM Executive Committee during the past 5 years. He has now been named president of Rockefeller University, effective July 1, 1978. He will be succeeded as SUMEX principal investigator by Professor Edward Feigenbaum, who is chairman of the Stanford Computer Science Department and has been closely associated with the resource since its inception. The coordination of project activities with medical research is the responsibility of Professor Stanley Cohen, Dr. Lederberg's successor as chairman of the Department of Genetics in the Stanford Medical School. Professor Lederberg will maintain close ties with these activities as chairman of the SUMEX-AIM Executive Committee and through his plans to encourage AI applications work at Rockefeller.

2) The SUMEX renewal application submitted last year at this time has been reviewed and approved by the National Advisory Research Resources Council.

Our proposed renewal term of 5 years was reduced to 3 years in view of the management changes in progress.

3) We have made a number of upgrades to the SUMEX facility hardware and software systems to enhance throughput and to better control the allocation of resources. We are also establishing a connection to the commercial TELENET network to explore more cost-effective ways to meet community communications needs.

4) We have made progress in the investigation of alternative schemes for the export of programs. A demonstration of the machine-independent MAINSAIL system is nearing completion for the initial set of target machines. The DEC 2020 system, formally announced early this year, provides a relatively inexpensive software-compatible machine for export or expansion of computing capacity for small research groups.

5) The progress of SUMEX-AIM user projects in the development of their respective programs is reported by the individual investigators. We have worked hard to meet their needs and are grateful for their expressed appreciation.

Valediction - Personal remarks by J. Lederberg

While Ed Feigenbaum and I cheerfully accept the full responsibility that is entailed by our roles as co-investigators of this resource project, we are embarrassingly aware how much of the effort has been the work of others. Choices for praise are always invidious, but I have no difficulty in singling out Tom Rindfleisch as the one person who deserves the most particular credit for the success of this program. His technical insight and finesse in the system design and implementation, and in the management of the resource staff are measured by the visible efficiencies and clarity of documentation of the resource. He is also preeminently responsible for the drafting of these reports and for managing our fortunes through all the complexities of federal and university accountability, and our obligations to local and national users.

Tom would be the first to insist on acknowledging the dedicated support of the administrative, programming and engineering staffs: I mention Carole Miller and Karen Carpenter, Rainer Schulz and Andy Sweer, and Nick Veizades as representatives of the several groups of veterans who have been part of SUMEX-AIM from its inception, and of the most conscientious team of my experience.

For my own role, I have leaned heavily on my friend and associate, Ed Feigenbaum, and it is gratifying to be so confident that the work we started together in building SUMEX-AIM will continue under his able stewardship. Elliott Levinthal and Bruce Buchanan did a great deal to make all this possible, and to make the tasks that Ed and I will have taken on not just manageable but fun. Carl Djerassi, in chemistry, was an indispensable fomenter of the scientific collaborations. Stan Cohen is making an equally great contribution, both by succeeding me as chairman of the genetics department, and by his continued promulgation of MYCIN and by serving as coordinator for medical school research interests in SUMEX.

But this list would eventually embrace a large part of Stanford University, a network of personal and interdisciplinary connections that constitutes a seamless web, a treasure for my own experience and recollection, -- but one that is perforce hard to fairly acknowledge, and even harder to sever myself from.

Fortunately, the communications net offers a way to soften that severance, and I will seek every opportunity to use it to stay in the closest contact with the affairs of SUMEX-AIM that the duties of my new situation allow. My continued association with Stanford and with SUMEX-AIM ought to be a self-exemplifying demonstration of the capabilities for community-building and for sustaining the human relationships in scientific effort that have been our highest hopes for these new, high technologies.

The realism of these expectations has been substantially tested already in the way that the Executive Committee of SUMEX-AIM, the user community, and Bill Baker and his colleagues at NIH/BRP, have been able to work together effectively and constructively in making this enterprise truly a national resource.

I look forward to continuing to be a part of a team like that!

2.1.2 TECHNICAL PROGRESS

The following material covers SUMEX-AIM resource activities over the past year in greater detail. These sections outline accomplishments in the context of the resource staff and the resource management. Details of the progress and plans for our external collaborator projects are presented in Section 4 beginning on page 61.

2.1.2.1 FACILITY HARDWARE

Over the past year, several significant changes have been made to the SUMEX hardware configuration and associated system software:

1) Core memory was doubled by adding 256K words
2) The file and tape system hardware was upgraded
3) A connection to TELENET is being implemented

The memory and file/tape upgrades have substantially improved system throughput and efficiency as discussed below. The TELENET connection is being established for evaluation as a possibly more cost-effective means for meeting community communication needs (see page 16). The current system hardware configuration is diagrammed in Figure 1 on page 9.

INTRODUCTION

The SUMEX-AIM facility has been operating at capacity in terms of prime-time computing throughput and user file space for the past 2 years as documented in our annual reports (see for example pp 4-8 of the 1976 report). This condition has constrained the growth of the AIM community and our ability to bring AI programs nearing operational status in contact with the potential external user communities while continuing to support on-going program development efforts. We have taken active steps to try to transfer prime time loading to evening and night hours including shifting personnel schedules (particularly for Stanford-based projects), to control the allocation of CPU resources between various user communities and projects, and to encourage jobs not requiring intimate user interaction to run during off hours by developing batch job facilities. Despite these efforts, our prime time loading has remained very high. Perhaps the most significant effect of the resulting poor response time is the deterrence of interactions with medical and other professional collaborators experimenting with available AI programs, whose schedules cannot be adjusted to meet computer loading patterns (see for example the MYCIN report in Section 4.2.6 on page 163).

Two years ago, the Executive Committee gave approval for the augmentation of SUMEX-AIM computing capacity by adding a second CPU. The decision for the CPU was made as a trade-off between adding memory and/or CPU to maximize capacity enhancement within the resources available (see the 1976 annual report for a discussion of these trade-offs). We implemented the dual processor system in the spring of 1976 and brought it on line in June. The additional capacity was put to use very quickly as reflected in system usage and loading data summarized in

Figure 6 through Figure 8. With the common criterion that users have pushed both the single and dual processor systems to the limits of useful work in terms of prime time responsiveness, it is clear that the second processor substantially increased throughput. The "tolerable" peak load average increased, the number of jobs on the system increased, and the number of delivered CPU hours increased. At the same time (as predicted) the overhead per machine rose dramatically as shown in Figure 9 and Figure 10. The overhead increases came principally in the category of I/O wait (total scheduler time and time waiting for a runnable job to be loaded in core) and in the time processing pager traps. Another factor, not explicitly shown in these data (because we only have a 1 msec clock), is the added time spent at interrupt level servicing drum swapping. This adds another 10-15% estimated overhead.

After the dual processor augmentation, SUMEX-AIM computing capacity again became overloaded. This continuing saturation has raised serious discussion about the scope of computing needs of the AIM community and possible justification of additional PDP-10 scale machines to be added to the AIM network. Several specific proposals have been submitted for additional user nodes. We expect additional capacity to be available through the Rutgers resource by the end of this summer and support expansions at other AIM nodes as justified by local and community needs. From the SUMEX viewpoint, we have attempted to do everything feasible and economically justified within current budgets to maximize the use of the existing hardware for productive work. After the dual processor augmentation, the obvious remaining CPU resource to be tapped was to reduce the high dual processor overhead.

A parallel saturation problem existed for a long time in file space commitments. We had queued requests from numerous projects for increases in file space including INTERNIST, Higher Mental Functions, Language Acquisition Modeling, DENDRAL, Chemical Synthesis, MYCIN, and several pilot projects. We did not have additional space to allocate to meet these needs and our DEC RP-10C controller configuration was full (7 drives on-line and one available for backup). We had taken an active role in trying to optimize use of available space by limiting the total space available to projects, limiting the number of versions of experimental files kept on the system, and encouraging the use of tape or Datacomputer archive services for files not needed on-line routinely. We still were unable to adequately provide for the growing needs of existing projects or meet the bare space needs of new projects getting started.

The following plan was presented to the Executive Committee to increase the capacity of the SUMEX facility configuration by 1) adding memory to optimize dual processor utilization and 2) redesigning the file system (including the tapes used to backup and archive user files) to meet increased demands within up-to-date technology. This plan was approved in June 1977 and implemented in September 1977.

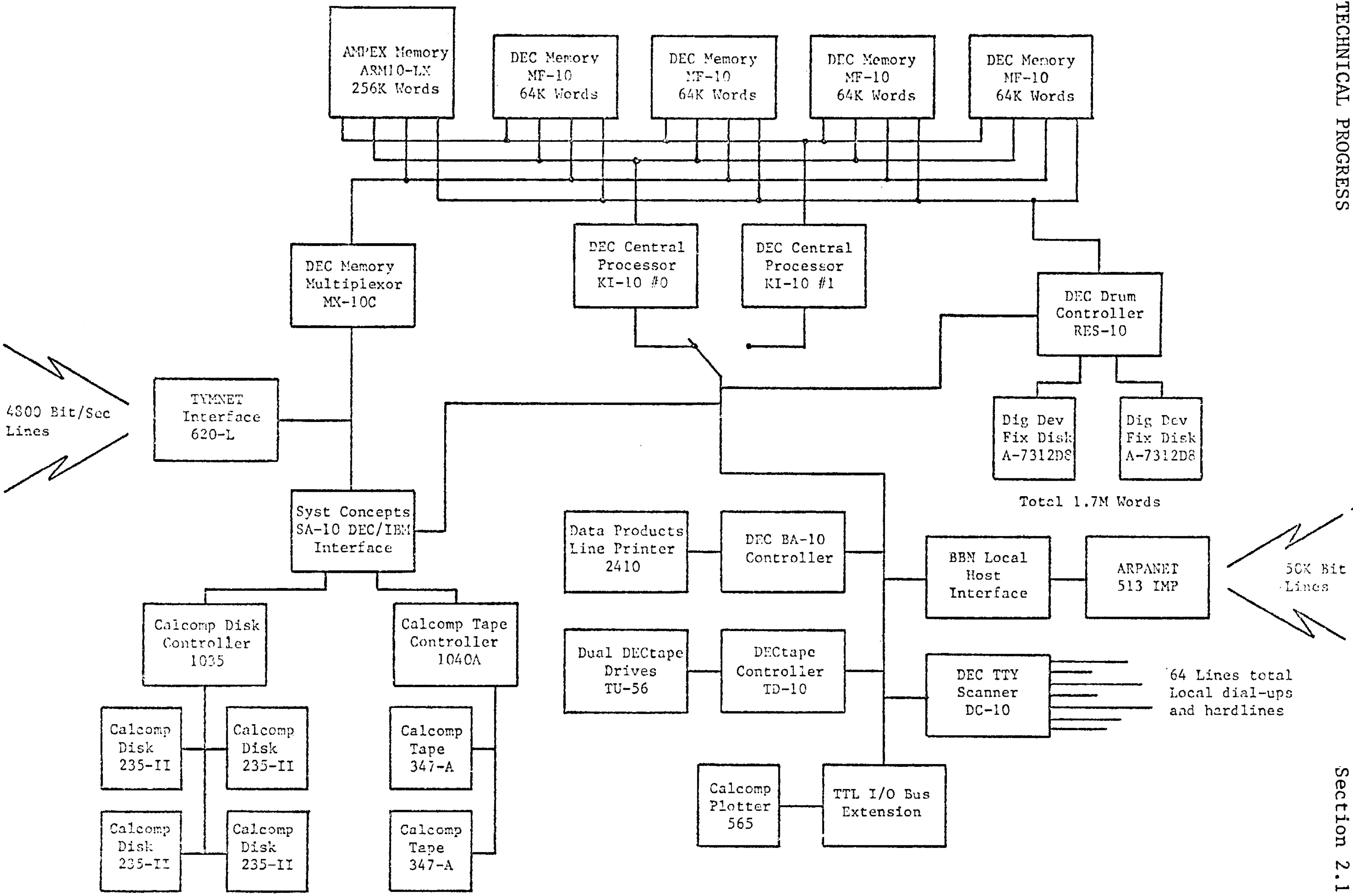

Figure 1. SUMEX-AIM COMPUTER CONFIGURATION (9/77)

MEMORY AUGMENTATION

There is a close interaction between memory size, CPU capacity, and secondary swapping storage in determining the performance of a demand-paged system like TENEX (see 1976 annual report pp 4-10). Our system as initially designed was quite well balanced in these respects. As the SUMEX-AIM computing load reached capacity, the choices for augmentation dictated either memory or CPU as we had insufficient funding to augment both. We chose CPU as the most user-effective means of providing more capacity at that time. However, as pointed out then, the added CPU power has the effect of increasing the system overhead in order to manage the increased number of jobs using the system within available memory. This shows up in increased pager trap time, interrupt-handling overhead for drum swaps, requirements for additional secondary swap space to accommodate the added jobs, and I/O wait time to fetch a runnable job into core as illustrated in Figure 9 and Figure 10.

Recorded data on dual processor performance show that, during prime time loading, the overhead in I/O wait time, pager trapping, and drum interrupt handling effectively amounted to about 40-50% of the second CPU (actually all of the I/O traffic is on one of the machines whereas other overhead for paging and I/O wait is distributed between them). This lost capacity was recoverable by adding memory. The effect of increasing memory is to allow more jobs in core at once (larger "balance set") and larger working sets to reduce overhead factors. This not only improves efficiency but smooths out user interaction since the larger balance set makes it more likely that pages for a given job will be in core when needed for teletype service. The 256K memory configuration only afforded a balance set of 4-5 jobs, so that with a load average of 7-10 during prime time, about 4-5 jobs had to be completely swapped out at a given time and hence could not get any service. The larger balance set means fewer jobs swapped out for a given load average.

The added memory size also allows more effective use of slower swapping space, particularly with the parallel disk system upgrade to the faster 3330 technology as outlined below. Having more jobs in the balance set makes it more likely that a runnable job is there and reduces the page fault rate so that swapping between memory and the slower store can occur without loss. Time previously wasted as I/O wait to get a disk-swapped job back into memory is reduced.

We considered a number of memory vendors and also looked into a new "slow" AMPEX memory which trades speed (3 usec versus 1 usec) for capacity (1000K words versus 256K words). This memory could be configured either in the form of a random access memory (RAM) or a block transfer ("drum-like") device. In essence the "block transfer" mode would add another layer in the hierarchy of storage intermediate between drum and high speed memory. We felt this type of "drum" memory would not be the most advantageous solution as the system was already burdened waiting for runnable jobs to execute and handling a high overhead of page swapping. The CPU time for the management of page swapping is non-trivial and, based on measurements of swapping activity, the overhead in managing storage on the KI-10 rivals the rotational latency of the drum so the fast swapper would not do that much good. The dominant factor in the overhead is that we had a relatively small executing store for our processing capacity so that under heavy loads the system thrashes trying to service runnable jobs for all the users.

Similarly, configuring such slow memory as RAM for our non-cache KI-10's would slow the processors down by a factor of 3 when executing code from the slow memory. It would be equally costly to rearrange pages between fast and slow memory. Without a special transfer device, the CPU would have to do the transfers limited by the slow memory speed. For a cache system (like the KL-10), this problem may be overcome by the firmware management of the movement of active memory locations between the very fast cache memory and the slower memory (4-word parallel transfers).

Thus we felt the most effective remedy was more high speed memory. We chose AMPEX memory from among the vendors reviewed as the best trade-off between performance, price, packaging, and maintainability. The additional 256K of memory was brought on-line in September 1977. From the data shown in Figure 9 and Figure 10, it is clear that the predicted reduction in system overhead was immediately achieved. The following table shows measurements of average instruction times comparing our AMPEX and DEC MF-10 memories. Also included for comparison are data for a Systems Concepts memory installed at the IMSSS KI-10 facility at Stanford:

	AMPEX	DEC MF-10	SC MF-10
MOVEI	1.20	1.21	0.97
MOVE	1.54	1.64	1.24

These data give the time in microseconds to execute the instructions shown based on the DEC timing diagnostic and normalizing the times to a "standard" 15 foot memory cable length. The MOVEI instruction shows the relationship between the basic memory access times and the MOVE instruction illustrates the effects of KI-10 "look ahead" with overall memory cycle time. Currently our AMPEX memory is timed to have essentially the same access time as the MF-10's but it is actually capable of somewhat faster operation. We are planning to attempt to reconfigure the memories this summer to take better advantage of the AMPEX speed. This may recover about 80 nanoseconds per access. This will still not bridge the timing difference between the AMPEX and the Systems Concepts memories. Systems Concepts offers a technically advantageous memory in terms of speed. Our choice of vendors was based on our own evaluation of issues like resale potential, maintainability, and management responsiveness, taking into account our experience with Systems Concepts in purchasing their disk channel interface (see below).

It should also be mentioned that the installation of the additional 256K of memory required modifications to the MX-10 memory multiplexor to accommodate 22-bit addresses and to the TYMBASE to be able to operate on a KI-10 style memory bus supporting more than 256K of memory.

DISK/TAPE RECONFIGURATION

Disk technology has changed rapidly in recent years. At the time we bought the initial SUMEX configuration, taking into account the discount DEC gave on the system purchase and maintainability, the DEC RP-03 system we bought was the best choice. Since then double-density 3330 technology has become well established (prices for IBM-compatible equipment were cut almost in half in 1976 alone!) and even higher densities are coming along. Given the relatively low incremental cost (for used equipment), we added RP-03 drives until filling the capacity of the controller. But with the added demands of community projects, a better long term solution necessitated upgrading from the RP-03 technology. Newer devices offer more economical future growth, and faster transfer rates thereby further decreasing system overhead.

Our tape system was in an even more advanced part of the age curve. We have not emphasized individual user tape services at all but tapes are critical to system operation for file system backup and user file archiving. We minimized the initial investment in tape drives to the advantage of other parts of the system. To accommodate the larger file system and to improve system operations and efficiency, the upgrade in file system also required a parallel upgrade in tape service. An additional advantage to upgrading the tape system was to move the I/O interface from the I/O bus to a direct memory interface thereby reducing system interrupt loading during prime time tape/file system operations.

The most attractive approach to file/tape system upgrade was to adapt a DEC memory port to look like an IBM selector or block multiplexer channel and then to take advantage of the substantial price competition in the IBM-compatible peripheral market. The capacity of a double-density 3330 disk drive is equal to 4 RP-03's. Thus bringing 3 new drives on line almost doubles the on-line capacity. After investigating alternative vendors, we selected a system using a System Concepts SA-10 channel adaptor, Calcomp 235-II disks, and Calcomp 347A tapes. This system was installed and brought on line at the same time as the memory augmentation in September 1977.

This system has substantially alleviated file capacity pressure and made possible much smoother backup operations. With the faster tape speed, we no longer take the system down for pack copies Sunday morning but rather do a full file system dump to tape. Similarly during the week we do incremental dumps back to the previous full dump each day to give quite good backup coverage. We have experienced no major technical problems with the new file/tape system; more details about impact on system software is given in a later section.

Unfortunately, we have experienced many frustrations dealing with Systems Concepts management in contrast to the high technical quality of their hardware. There remain several parts of the SA-10 adaptor that have not been delivered including full documentation, maintenance training, cabling to replace that which we borrowed for installation, and the device indicator panel. This experience led us not to consider Systems Concepts for memory. Few such memories have been delivered and it is not clear that we could depend on future maintainability. On the other hand, support for the SA-10 is secure in that many are in the field with excellent service records and alternative sources exist for SA-10 maintenance through Calcomp, DEC, or TYMSHARE.

2.1.2.2 SYSTEM SOFTWARE

MEMORY EXPANSION AND FILE/TAPE UPGRADES

The addition of 256K of memory and the upgrades of our file/tape system necessitated a number of changes in the monitor. TENEX had not fully anticipated memory addresses longer than 18 bits and so those places where half-word addresses were assumed had to be fixed. The RP-10C disk service and TM-10 tape service code had to be replaced by code that produces the appropriate IBM channel commands. We imported the "standard" BBN SA-10 disk/tape service and incorporated it in our dual processor system. Despite the substantial amount of work required to incorporate this code into our system, the new hardware and monitor came up smoothly on 9/1/77.

We have encountered a number of bugs during the year, particularly in the disk service. The most troublesome one resulted in a deadlock between command retry attempts from the CPU and a busy controller state. The 3330 recovery procedure implemented in the BBN code appeared to track exactly that used in IBM's most recent VS releases. Nevertheless infrequently during internal controller error correction attempts, we found the system hung in a loop with the controller, when trying to restart commands queued at the time of the error. It appears the problem may be in the Calcomp controller microcode but we have not been able to get enough information from Calcomp to confirm that. Meanwhile we have constructed a software work-around to detect the loop when it occurs and to reset the disk channel before proceeding.

The new hardware had other ramifications for system software as well. DEC's diagnostic system is designed to run off of their disk or magnetic tape systems. This capability was lost as a result of the change from DEC hardware so that diagnostics had to be loaded from slow DECtape units. We have invested considerable effort in bringing system diagnostic facilities back up to a workable level; borrowing programs others had written and implementing new ones where needed. We have implemented a stand-alone facility to load SAV files from the TENEX file system, incorporating full TENEX name recognition features. This means that programs can be manipulated and kept on-line in the time-sharing file system and then loaded as needed when the machine is down or in stand-alone mode. This also provides an easier way to reload the monitor. We have written a fast disk pack copy routine for the 3330 packs and have improved the SA-10 diagnostic package to check out 512K of memory and to ensure safer testing of disk drives in the presence of live file system packs.

Also with 1600 BPI tapes available changes were necessary to the tape service and TOPS-10 compatibility package to accommodate DEC's extended magnetic tape UUO's as well as to be able to fully use the byte packing facilities of the SA-10 and IBM drives.

We continue to work to improve the efficiency of the system and its effectiveness in allocating valuable resources. We have implemented a high priority hardware clock to sample monitor and user mode program counter locations to find places of abnormally high activity and perhaps inefficiency. This has pointed out several "hot spots" in routines where it was obvious on other grounds that a substantial amount of time is spent (e.g., drum service, KI page handling and teletype service) but there are no clear solutions to these problems with the current hardware.

SYSTEM LOADING CONTROLS

We previously implemented a form of "soft" CPU allocation control in the monitor, assisted by a program which adjusts user percentages for the scheduler based on the dynamic loading of the system. The allocation control structure works based on the scheduler queue system and takes account of the *a priori* allocation of CPU time and that actually consumed. Our TENEX uses a hierarchy of five queues for jobs ranging from highly interactive jobs requiring only small amounts of CPU time between waits to more CPU intensive jobs which can run for long periods without user interaction. These interactive queues (text editing, etc.) are scheduled at highest priority without consideration of allocation percentages. If nothing is runnable from the high priority queues, the CPU-bound queues are scanned and jobs are selected for running based on how much of their allocated time has been consumed during a given allocation control cycle time (currently 100 seconds). If no such jobs are runnable, then those that have received their allocation of CPU time already are scheduled based on how much they are over allocation and how long they have waited to be run again. This system is not a reservation system in that it does not guarantee a given user some percentage of the system. It allocates cycles preferentially, trading off *a priori* allocations with actual demand but does not waste cycles.

This scheduling scheme does not deal with the problems of system overloading during peak periods. At such times (mid-morning and mid-afternoon especially), one observes what has been termed "the tyranny of time-sharing". System efficiency and user response time degrade because the system is trying to serve more jobs than it has reasonable resources for. Users sit at their terminals waiting for the cycles they need to work effectively but there are not enough to go around. Ideally the system should have a response time keyed to a typical human interactive response interval. This implies a limit on the number of active job slots that the system can accommodate simultaneously in order to approximate this ideal. In some systems a "pie-slice" scheduler is used wherein a group of users is allocated some percentage of the machine and if that group consumes more than that amount during the cycle interval, its jobs are not scheduled until other groups catch up. Meantime, those users sit at their terminals and receive VERY slow response, not knowing when things will let up so they can effectively compute again. This type of approach does keep the system from trying to run too many jobs at once but it does not solve the problem of EFFECTIVELY MANAGING USERS' TIME.

We have attempted to control system overloading in a somewhat different way to better manage user time and to allow us to better apportion system capacity between communities and projects during heavy load. Each project gets its *pro rata* share of the active job slots the system can accommodate. Rather than allow many users to ineffectively vie for each project's slots (as in the pie-slice system), we ask selected users within each group to restrict their use for periods of 1/2 hour so that those remaining can work effectively within the project aliquot. Allocation of active job slots is made on the basis of relative community and project percentage allocations (assigned by the AIM Executive committee). Within each project slots are allocated either on a round-robin basis or taking into account optional project priorities among users. Under overload conditions, active jobs outside of the available slots are asked to slow down, thereby holding the load within tolerable limits. If such jobs do not voluntarily cooperate, they may be forced to comply.

An overload condition is defined to be one in which the overall load average exceeds a threshold (currently 7.5), significantly more jobs are runnable than there is core for, or excessive page faulting is occurring. Outside of periods of overload, the previous "soft" percentage scheduling scheme is applied. Also thresholds for overload conditions may be dynamically adjusted to assure good system response during a demo.

This system has been in operation for approximately one month during which time we have experimented with various threshold adjustments and observed its effects on user behavior. During this early period, we have not placed any controls on system use by the AIM community projects since they have historically been below their quota for system use. It is still early to tell quantitatively what its effect will be; system usage fluctuations are such that we will have to observe operations for several months before drawing conclusions. However, qualitatively it seems to be holding system loading within tolerable bounds and allocating capacity as apportioned between the various communities and projects.

OTHER ENHANCEMENTS

Other areas of system software development include the EXECutive program, the BSYS program for file archiving and retrieving, the printer spoolers, the CHECKDSK program for verifying file system integrity, and numerous smaller utility extensions and bug fixes. We have continued to improve the EXEC in such areas as the DIRECTORY command (to display last file reader and temporary files), the MAP command (to handle long files), a new INITIALIZE command (to restart the EXEC after errors), smoothing out multi-directory search paths using features of the new GTJFN, adding wild card and question mark facilities to the file retrieval INTERROGATE command, making the command for changing file protection more mnemonic, and restoring terminal modes correctly after a forced detach (e.g., with a dropped line or network disconnect).

We have added a facility to the BSYS program to automate the restoration of requested files from tape, avoiding the earlier error-prone and time-consuming typing of individual restore commands. We have completely rewritten the line printer spoolers to more efficiently and uniformly handle the local and remote printers, to add facilities for "unlisting" a file listed by mistake, and improving the marking of listing boundaries to ease operator separation of listings for various users.

We have imported a new version of CHECKDSK initially written at BBN and have incorporated local facilities for more extensive file system integrity checking. This version presorts file index block addresses and scans for errors using more sequential disk I/O. Several forks are started, one for each drive. These keep the disk channel as busy as possible while performing the check computations. This improvement has reduced the time to scan the file system from 20 to 6.5 minutes.

2.1.2.3 NETWORK COMMUNICATION FACILITIES

A highly important aspect of the SUMEX system is effective communication with remote users. In addition to the economic arguments for terminal access, networking offers other advantages for shared computing. These include improved inter-user communications, more effective software sharing, uniform user access to multiple machines and special purpose resources, convenient file transfers, more effective backup, and co-processing between remote machines. Until now, we have based our remote communication services on two networks - TYMNET and ARPANET. These were the only networks existing at the start of the project which allowed foreign host access. Other commercial network systems (notably TELENET) have come into existence and are growing in coverage and services.

Users asked to accept a remote computer as if it were next door will use a local telephone call to the computer as a standard of comparison. Current network terminal facilities do not quite accomplish the illusion of a local call. Data loss is not a problem in network communications - in fact with the more extensive error checking schemes, data integrity is higher than for a long distance phone link. On the other hand, networking relies upon shared community use of telephone lines to procure widespread geographical coverage at substantially reduced cost. However, unless enough total line capacity is provided to meet peak loads, substantial queueing and traffic jams result in the loss of terminal responsiveness. Limited responsiveness for character-oriented TENEX interactions continues to be a problem for network users.

TYMNET:

Networks such as TYMNET are a complex interconnection of nodes and lines spanning the country (see Figure 2 on page 18). The primary cause of delay in passing a message through the network is the time to transfer a message from node to node and the scheduling of this traffic over multiplexed lines. This latter effect only becomes important in heavily loaded situations; the former is always present. Clearly from the user viewpoint, the best situation is to have as few nodes as possible between him and the host - this means many interconnecting lines through the network and correspondingly higher costs for the network manager. TENEX in some ways emphasizes this conflict more than other time-sharing systems because of the highly interactive nature of terminal handling (e.g., command and file name recognition and non-printing program commands as in text editors or INTERLISP). In such instances, individual characters must be seen by the host machine to determine the proper echo response in contrast to other systems where only "line at a time" commands are allowed. We have seen little improvement in TYMNET service over the past year although the cost of service has risen sharply. We purchase TYMNET services through a volume contract the National Library of Medicine has with TYMNET. The cost has gone from approximately $2.90 to $6.09 per connect hour. Because of this increase, we are investigating alternative sources of network service; in particular TELENET.

We have had a number of technical problems with the TYMNET this past year. Internally they changed some of the protocol involved in the TYMBASE connection we use. They neglected to tell us about these changes though and the problems that resulted were very hard to track down. From the user viewpoint, connections were dropped frequently. From the system viewpoint, we could not tell if the problems were subtle results of the recent memory and file system hardware change. That change required a modification in the TYMBASE to accommodate the new memory bus conventions. This took months to isolate but the TYMNET interface is finally running reliably again after much user frustration.

ARPANET:

Current ARPANET geographical and logical maps are shown in Figure 3 and Figure 4 on page 19. Consistent with agreements with ARPA and the Defense Communication Agency, we are enforcing a policy that restricts the use of ARPANET to users who have affiliations with ARPA-supported contractors and system/software interchange with cooperating TENEX sites. We have maintained good working relationships with other sites on the ARPANET for system backup and software interchange. Such day-to-day working interactions with remote facilities would not be possible without the integrated file transfer, communication, and terminal handling capabilities unique to the ARPANET.

TELENET

We recognize the importance of effective communication facilities for SUMEX-AIM users and are continuously looking for ways to improve our existing facilities. A year ago we did some preliminary investigations of TELENET facilities that have been rapidly expanding this past year (see Figure 5 on page 21). BB&N has hooked one of their TENEX systems up to TELENET and whereas we did not have the same quantitative tools we have for measuring response on the TYMNET, we observed TELENET delays at least as long as those encountered on TYMNET. We did the reverse experiment by using long distance telephone to connect from the TELENET node in Washington, D.C. to the SUMEX machine in California and observed the same sort of delays reaching several seconds per character. The TELENET has many attractive features in terms of a symmetry analogous to that of the ARPANET for terminal traffic and file transfers and being a commercial network, it does not have the access restrictions of the ARPANET. As indicated above, the cost of using the TYMNET has increased this past year so that TELENET rates appear to be substantially lower for supporting community communication services. The National Library of Medicine has a contract with TELENET which includes significant cost advantages through combining our use with NLM use to achieve a high volume discount. As a result of discussions with the AIM Executive Committee and BRP, we are in the process of implementing an experimental connection to TELENET with the view of moving SUMEX users to that service if technically effective.

Figure 2. TYMNET Network Map

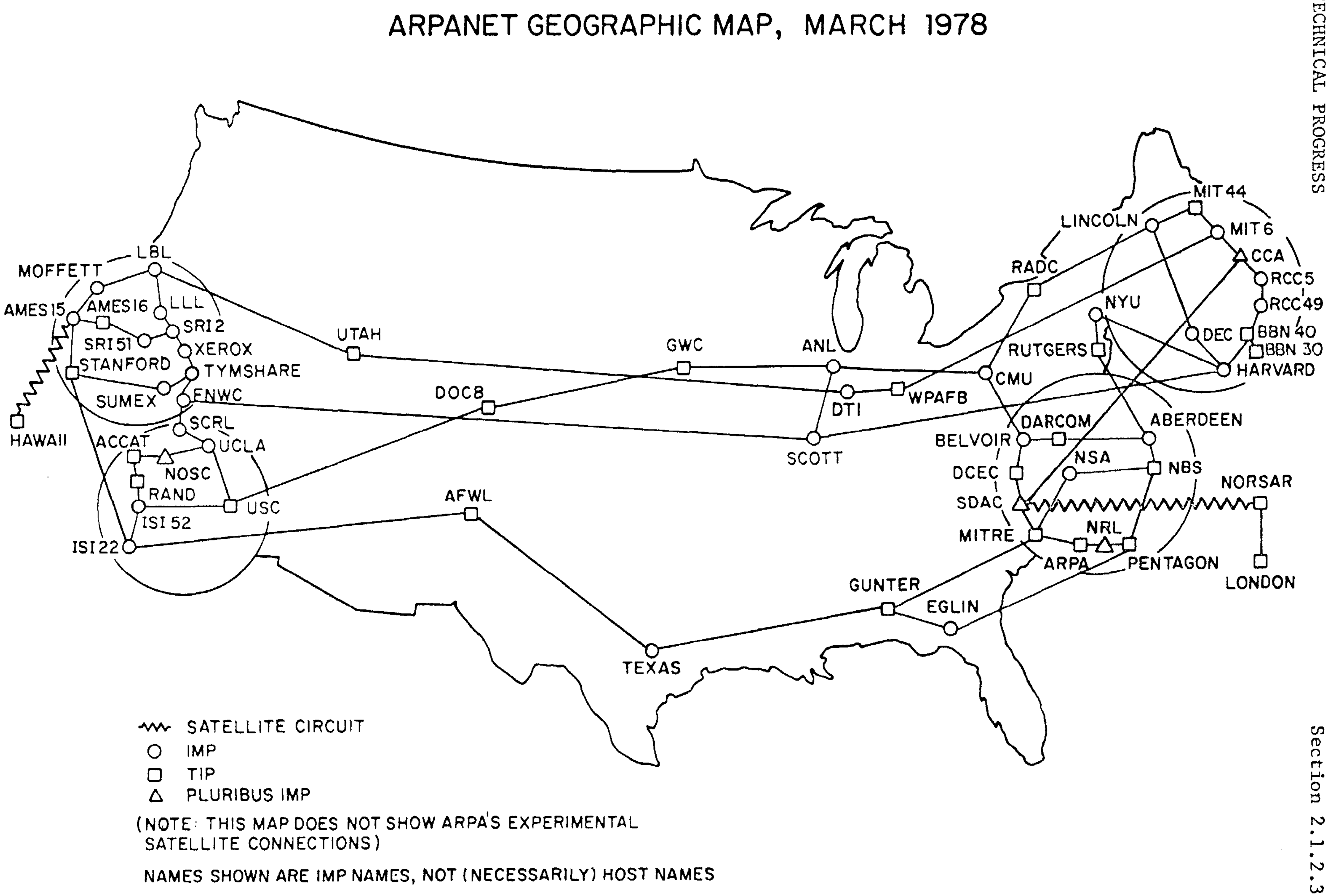

Figure 3. ARPANET Geographical Network Map

ARPANET LOGICAL MAP, MARCH 1978

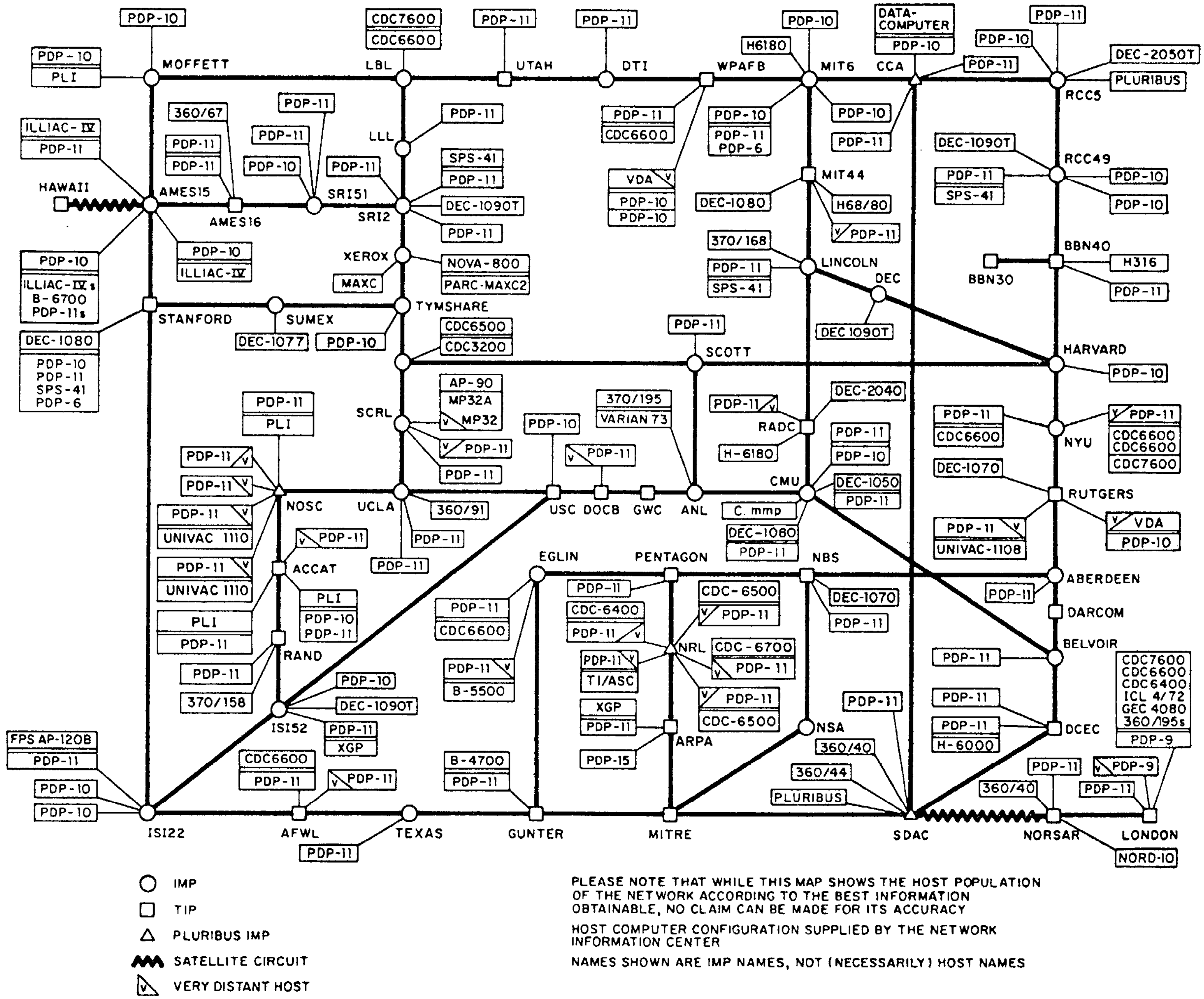

Figure 4. ARPANET Logical Network Map

TELENET GEOGRAPHIC MAP

MID 1977

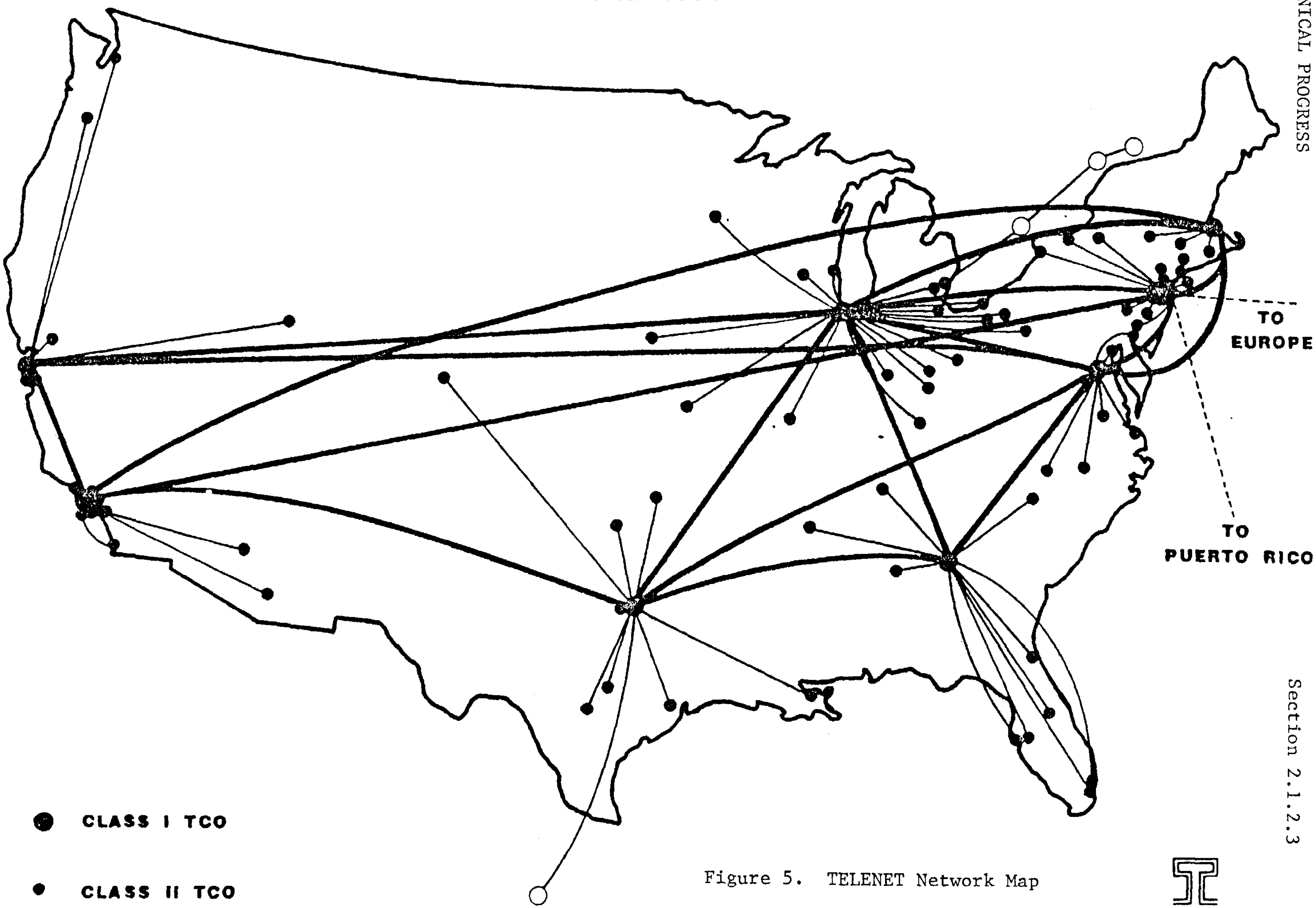

Figure 5. TELENET Network Map

2.1.2.4 SYSTEM RELIABILITY AND BACKUP

System reliability has been very good after the installation of the new memory and file/tape hardware. There have been a number of problems as detailed earlier with the disk system and TYMNET that have caused more crashes and dropped lines than normal. Also in the process of experimenting with speeding up our memory configuration, we have caused some unreliability. The table below shows monthly downtimes for the past year.

	1977								1978			
	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR
CRASHES												
Hardware	12	27	8	18	6	6	14	6	6	5	2	4
Software	0	0	4	4	2	1	3	4	1	0	6	3
Environmental	1	2	2	2	2	0	0	1	1	1	1	4
Unknown Cause	0	2	3	2	3	5	6	5	4	7	3	2
DOWNTIME (Hrs)												
Unscheduled	30	70	34	47	19	19	48	20	17	25	13	15
Scheduled	41	73	25	83	61	44	31	30	23	15	28	31

TABLE 1. System Reliability by Month

In May and June we experienced a substantially higher number of hardware crashes which we feel resulted from the system being overheated during an air conditioning failure in mid-May. Ultimately an intermittently shorting backplane wire was found as well as several intermittent arithmetic unit failures. During late July and August we worked on system hardware and software changes preparatory to the installation of the additional 256K of memory and the new file/tape system hardware. This increased system downtime and unreliability as well. Infant mortality problems with the new hardware (AMPEX memory especially), caused a number of crashes in August, September, and October. The high number of "Unknown" crashes until recently are the result of a number of factors which were hard to separate or caused the system to fail in ways that we couldn't reconstruct what happened. We feel these resulted from the TYMNET protocol problem mentioned earlier, a race condition between the AMPEX and DEC memories, disk controller problems, and software bugs. Between September and late October we worked on organizing and enhancing system diagnostic capabilities to support the new hardware. This required increased downtime.

2.1.2.5 SOFTWARE EXPORT

Over the past year we have continued to investigate alternatives for software export. The following reports on several of these areas, including 1) the availability of small PDP10-like machines, 2) progress in developing the MAINSAIL language, and 3) an investigation of possibilities for writing a MYCIN-like program using an algorithmic language.

SMALL PDP10-LIKE MACHINES

Early this calendar year DEC announced a new small machine designated the 2020. This machine approximates the "small PDP-10" machine we had discussed in last year's report. It allows up to 512K of memory, 2 RP-06 disk drive, 2 tape drives, a printer, and runs DEC's TOPS-20 operating system. Prices range from $150K for a 128K word machine to $375K for a 512K word machine with two disk drive, tapes, and a printer. The unloaded performance of this machine appears to be in the range of a KA-10 but only preliminary benchmarks have been run. Lynch at SRI has run a simple LISP test program on a range of machines. His test creates a large list, randomizes it, and then sorts it. Using a KA-10 with 512K of memory as a reference, the performance of various PDP-10 systems is shown in the following table.

Function	KA-10 256K	KA-10 512K	2020 512K	KI-10 512K	KL-1090T 1024K
Build List	0.94	1.00	1.20	1.79	5.86
Rearrange List	0.92	1.00	1.13	1.64	5.20
Sort List	0.79	1.00	0.89	1.65	4.56

This test indicates that the 2020 performs at about the same level as the KA-10 for a single user. These data do not give a complete picture of performance under increasing load, however, and do not fully reflect the intrinsically slow arithmetic performance of the 2020.

We have attempted to run benchmarks of the CONGEN and MYCIN programs to compare these machines. We ran squarely into a compatibility problem however. We prepared the two benchmarks, ran them on our KI-TENEX system, checked to see that they would run on TOPS-20 using SRI's 1090T system, and asked DEC to run them on the 2020. The benchmarks failed to execute because of some system call changes DEC had just made to the newest release of TOPS-20 running on the 2020. We are just now getting access to a machine running that version of TOPS-20 and hope to fix the incompatibility to complete the benchmarks. This experience reinforces our belief that increasing incompatibilities will show up between TENEX and TOPS-20 that will make software transfer difficult.

We have had a number of contacts from outside users interested particularly in the chemistry AI programs. Such a machine would represent a good solution for such groups to gain access to the programs, maintain the necessary security for proprietary data storage, and stay abreast of new developments. This type of approach is also attractive for providing needed capacity expansion in small increments within the AIM community (e.g., for more extensive testing of the MYCIN or INTERNIST programs) while maintaining general software compatibility. Remote location of such machines within the community may also offer significant advantages for human interfaces since terminal handling can be done locally thereby supporting higher speed lines and improved echo interactions for recognition, etc.

MAINSAIL

During this past year we have concentrated on six areas:

1) Implementations
2) Runtime design
3) Language design
4) Compiler design
5) Documentation
6) Emulation research

We have not yet extensively distributed MAINSAIL since it is still undergoing development based on our experiences with it locally. We have continued to receive many inquiries concerning the progress of our work, with several projects considering using MAINSAIL when it becomes available.

At present our major concern is MAINSAIL's efficiency in a small address space; in particular, the compiler cannot yet run on a PDP-11. Though it appears that computer technology is moving towards large address spaces, existing machines with 32K-word address spaces will persist for many years, and many people have indicated an interest in using MAINSAIL on such machines. The difficulty is that MAINSAIL provides features which are not easily supported when memory is scarce. Over the past year we have gained a better understanding of MAINSAIL's resource requirements, and have taken steps to reduce its implicit use of memory.

Implementations

We have developed five implementations for two computers: TOPS-10 and TENEX for the PDP-10; and RT-11, RSX-11M and UNIX for the PDP-11. The last two were developed during the past year. The others have received varying amounts of bug fixes and updates. The TENEX version has been in use for about two years, and the RT-11 version for about a year. The TOPS-10 version has been used to a lesser extent for about a year. Programs have been run on RSX-11M and UNIX, but these implementations are not complete.

No implementation is in general use; in some cases they have primarily served to insure that the runtime design is sufficiently flexible. Each new implementation has revealed deficiencies in the design which have since been corrected. We will need to implement MAINSAIL on some non-DEC machines before we can get an unbiased assessment of the difficulty of creating new implementations.

Runtime Design

A new runtime system is now under implementation. It is oriented towards execution efficiency and less memory utilization since these are the problems with the current PDP-11 implementations.

A major savings has been made with regard to string constants. In the previous implementation, the text of string constants was copied into string space where it remained throughout execution. Also, the string-constant descriptors were allocated in the data sections, which remained in memory as well. In the new implementation, the string-constant text remains in the control section, so that it is swapped out of memory along with the control section. String-constant descriptors are created each time a string constant is used. This usually requires that the text for a string constant be copied into string space upon each use. The overall result is that string constants do not tie up memory as in the previous implementation, but more time may be spent repeatedly copying string constant text into string space. This could also lead to more string "garbage" collections.

The implementation of modules in terms of control sections, data sections, and descriptor sections has been altered to save memory. Procedure call, entry, exit and return have been redesigned to save code and time. The amount of code executed for i/o has been decreased. A new approach to the use of "anonymous" modules has been implemented, and the manner in which modules obtain linkage to one another has changed.

The previous implementation required that every module reside in a separate file which is opened and closed during execution in order to access the module's code. The new implementation provides "runtime libraries" which are files containing any number of modules. Each runtime library remains open throughout execution. There will be a standard runtime library containing the system modules, and another containing the compiler modules. The programmer may also contribute runtime libraries. Single-module files remain as before.

The size of the "kernel" module (which is always resident) has been decreased. There is more reliance on incremental initialization of arrays, string space, string constants, class descriptors, module pointers and module descriptors. This allows some code to be moved out of the kernel into separate modules.

The modules which make up the runtime system have been reorganized to decrease the number of costly intermodule calls. Each module is relatively more self-contained. In the previous implementation, many calls to system procedures resulted in a chain of intermodule references which resulted in thrashing on the PDP-11. In most cases a call to a system procedure now requires at most a single system module.

A preliminary version of a debugging module has been written and utilized to some extent.

Language Design

There have been some changes to the language, primarily to support the new runtime implementation.

OWN arrays are no longer handled any differently than other OWN variables. An OWN array's allocation is now under programmer control. An OWN array's declaration may no longer include initialization values. Instead, an INIT statement is provided which can initialize any array with constant values.

To reduce the number of intermodule calls, and thus the amount of potential swapping and the extra code executed for the calls, the concept of "compiletime libraries" has been introduced. A compiletime library is a file containing procedure bodies that are to be "compiled into" a number of different modules instead of being interface procedures of a separate module. The use of new compiler directives, and an extended concept of FORWARD procedures, support the use of compiletime libraries. Small system procedures have been put into a standard compiletime library which is automatically utilized during the compilation of any module which invokes one of the procedures.

cmdFile (command file) and logFile (logging file) are now the standard input and output files used by the compiler and the runtime system where tty (i.e., the user's terminal) was previously used. Both are initially opened to tty. These files are used instead of tty so that the user can "redirect" the standard input or output stream if desired. This can be done via subcommands at the start of program execution, or via explicit opens during program execution. The system procedures ttyRead and ttyWrite still provide direct communication with tty.

The CHECK and NOCHECK compiler directives have been implemented. CHECK directs the compiler to henceforth emit code to check certain conditions (such as array subscripts and NULLPOINTER's) at runtime which cannot be determined at compiletime. NOCHECK can be used to turn such checking off. These directives have been in the language (in a slightly different form), but they had never been implemented in the code generators.

MAINSAIL has previously guaranteed ASCII character codes. This requires a translation on machines with other character codes (e.g., the IBM-370 uses EBCDIC). Our experience shows it is not difficult to write programs independently of the character codes if certain minimal assumptions are in effect. For example, the characters A...Z are guaranteed to be in alphabetic order, but they are not necessarily contiguous. We have also introduced eol (end-of-line) in place of the previous ASCII-dependent crlf (carriage-return-line-feed), and eop (end-of-page) in place of the ASCII form-feed. eol and eop are defined as implementation-dependent string constants. Character incompatibilities among machines is a difficult problem to deal with, and may ultimately complicate MAINSAIL's implementation on machines with "deficient" character sets (e.g., a CDC with a 6-bit character set).

Compiler Design

Two steps have been taken with the goal of getting the compiler to execute in a small address space. First, it has been broken into smaller modules. It used to consist of about 10 modules, but now consists of about 60. This allows a more accurate working set of modules to build up in memory since no one module is so large that it displaces most of those currently resident. However, there is more overhead involved in initializing so many modules on machines which have sufficient memory for the entire compiler. The second step has been the size reduction or elimination of many of the compiler's data structures. Where possible data is maintained on a file rather than in memory.

To save the space required by the text of string constants, the compiler error messages have been placed on a file. Calls to error message procedures specify the location on the file of the appropriate error message. A program has been written to generate a new compiler error message file by combining the messages on an existing error message file with any new messages specified as string constant arguments to error procedures in the compiler modules.

The code generators have been modified to output additional information (in a separate file) for debugging use during execution. This debugging file is created only if requested as a subcommand to the compiler. It contains a symbol table and a cross reference between the generated code and the source files.

Documentation

A new MAINSAIL manual is nearly complete. The former manual was an alphabetically ordered reference manual written primarily for internal use, that is, for those who were either already familiar with MAINSAIL or who at least knew SAIL. The new manual is a more readable reorganization and expansion of the reference manual information, organized by topic (e.g., data types, procedures, modules), and incorporating numerous examples.

Concurrent with the writing of the manual was the development of a formatting program to input the manual as written (with encoded section numbers and index references, no table of contents, etc.) and output a complete, ready-to-print manual.

An invited paper entitled "The MAINSAIL Project: Developing Tools for Software Portability" was delivered at the First Annual Symposium on Computer Application in Medical Care, given in Washington, D.C. in October, 1977. This has resulted in a number of inquiries from researchers interested in MAINSAIL's portability.

Emulation Research

The goal of the emulation research is to determine efficient means of representing MAINSAIL programs for interpretive execution. If the interpreter can be written in the microcode of the host machine, the resulting emulation should be more efficient than execution of MAINSAIL translated into a standard machine code. This approach simplifies compilation, and allows the efficient monitoring of program execution, so that debugging capabilities and performance measurements can become an integral part of program execution.

Statistics gathered from programs written in MAINSAIL have been used to guide and justify the design of a language representation suitable for emulation. An interpreter has been developed, and the static and dynamic properties of the resulting high level MAINSAIL interpretation are now under study. Its properties will be compared with conventional machine language implementations of MAINSAIL. The generated code appears to be about a third the size of standard machine code. It is more difficult to measure execution time differences since the processor design must be taken into account.

The characteristics of a suitable host processor to support the emulation are being examined in detail. It appears that a "universal host" (i.e., a processor not designed with a particular representation in mind) will not be able to execute a tailor-made representation as fast as a conventional processor can execute a standard machine language representation. Thus a "poor" representation on a processor designed to execute that representation seems faster than a "good" representation on a processor not designed for that representation. For this reason a microcode and processor organization which are oriented toward execution of the MAINSAIL representation are under design.

ALGORITHMIC LANGUAGE IMPLEMENTATIONS OF MYCIN-LIKE SYSTEMS

Production systems (PS) have been used extensively for knowledge representations for a number of AI applications such as MYCIN (2). Traditionally these systems have been implemented in various dialects of LISP. This has been so partly because LISP contains several "natural" representations for PS's, and partly because of the unique development and debugging environment offered by systems like INTERLISP.

Because of its generality and power, however, INTERLISP tends to be a expensive system to run, requiring a large amount of computational time and a large address space. While the expense of such a system is justifiable in a research and development environment, it may not be in more operational environments where these programs are to be used. One way to overcome this cost is by designing more economical LISP systems - several groups are working on this. These systems are typically built around special-purpose "LISP machines". Over the past year we examined another alternative, converting a MYCIN-like system into an algorithmic programing language (such as SAIL or MAINSAIL). This approach may offer advantages in being able to run versions of MYCIN-like systems on existing laboratory computers. The objectives of this study were to find ways to trim the resource requirements of the system while preserving as much of the knowledge representation clarity and modularity of the production system approach as possible. One of the major hurdles to be overcome is the difference in program and data representations between LISP and algorithmic languages. Several different approaches and languages were explored. Some are implementable in almost any algorithmic language, while others exploit features unique to certain languages.

The principal design features of MYCIN include (3):

1) A rule based consultation system. The knowledge is represented as collection of production rules. The consultation is driven by a goal directed search of the knowledge base.

2) An examination program which will explain the "line of reasoning" the system has gone through to produce the current consultation.

3) A question-answer system to query the system on parts of the consultation, or to ask general questions of the knowledge base.

4) A method of updating the knowledge base, by adding new rules or changing or deleting incorrect rules.

The knowledge base for the system is stored as a collection of production rules in the form of PREMISE-ACTION pairs. The consultation is driven by a goal-

---

(2) See for example, Davis, R., Buchanan, B., and Shortliffe, E., "Production Rules as a Representation for a Knowledge-Based Consultation Program," Artificial Intelligence, Vol 8, No 1, February 1977.

(3) see for example, Shortliffe, E.H., "Computer-Based Medical Consultations: MYCIN," Artificial Intelligence Series 2, Elsevier, New York, 1976.

directed search of these rules, i.e. if the PREMISE of a rule depends on the value of a given parameter and its value is not known, rules which conclude something about that parameter value are evaluated. The flow of control depends on an interaction between goals stated in the IF clause of one rule, and the THEN clause in others where this interaction changes as the rule base is changed. This can be readily achieved in LISP by storing the executable PREMISE-ACTION routines as properties of the various rules and linking them in a list. This unity between program and data is one of the key points in the difference between a LISP approach and one using an algorithmic languages.

Considerations in Non-LISP Approaches to MYCIN

MYCIN can be viewed as a very complex "IF THEN" clause, but this ignores the flexibility and modularity of knowledge representation of a PS. At the very least, the programming language should have the capability of creating LISP-like data structures, such as trees and lists. If we examine the nature of the rule interpretation, we can see that it is recursive in nature, so the language should also support recursive procedures. Many modern languages such as SAIL, PASCAL and MAINSAIL have these features.

The main problem is to represent the rules in such a way that they can be executed or interpreted in some sense, but can also be woven into the data base so that they may be fetched, examined, and modified when needed. We need to unify the control store and the data store in a way similar to that of LISP. We have looked at two approaches: one in which the rules are represented strictly as data, and interpreted as needed, and one in which the rules are represented as procedures.

Rule Interpretation Approach

In this approach the production rules are represented strictly as data, and procedures written to interpret the data structure. This is equivalent to writing a small, special purpose LISP interpreter. Any language which fills the requirements outlined above can be used to write the interpreter. The general procedures needed for this approach are:

1) A procedure to insure that atoms are unique

2) Procedures to read and write the data structures.

3) A procedure for each of the LISP functions to be executed, e.g., logical operations.

4) A procedure equivalent to LISP's EVAL, which will examine a list and invoke the correct procedures to interpret it.

This scheme has many of the advantages of the LISP version. It is easy to add new rules to the system that use only the currently defined functions. The hope would be that this interpreter would be smaller and faster than equivalent more general LISP machinery. Some disadvantages are that if a new function is needed, the EVAL section of the program must be rewritten and recompiled. The effort involved on this can be minimized by proper modularization, however.

Procedural Approaches

In these approaches we generate procedures which perform the following functions for each rule in the knowledge base:

1) Execute the PREMISE of the rule, and return a value indicating whether or not the PREMISE is true. In addition there must be a mechanism for marking which clause of the premise failed, if any.

2) Execute the ACTION of the rule.

3) Return a list of the parameters appearing in the PREMISE of the rule (also marking which clause they occur in).

4) Return the parameter (or list of parameters) referenced in the ACTION of the rule.

5) Print an English version of the rule.

These procedures can be generated automatically either translating LISP rules from MYCIN or translating an English input. In this way, we have replaced the interpretation of the rules as data structures with procedures that return truth values and carry out the actions. Besides the procedures, there is also a data structure which represents the interconnection of the rules. This approach could be faster than the interpreter for rules since we effectively have interpreted the rules once and for all at compile time.

Case Statement Procedures

In this implementation, a procedure is constructed for each of the functions mentioned above. The procedure takes an integer representing the rule number, and executes the proper subsection for that rule. For example, we might have:

```
BOOLEAN PROCEDURE premise ( INTEGER ruleNo );
        CASE ruleNo OF
                BEGIN
                [ 1 ]   < PREMISE of RULE001 >
                [ 2 ]   < PREMISE of RULE002 >
                          .
                          .
                          .
                END;
```

By marking a parallel data structure, we can trace which rules and which clauses of which rules have been executed.

A disadvantage to this method is the relative inflexibility of the rule base once it has been written. This can be alleviated to some degree by proper modularization.

MAINSAIL Implementation

This approach exploits some of the unique features of MAINSAIL. A MAINSAIL program is broken up into a number of MODULE's, which communicate with each other by means of interface fields. Each rule is represented as a different module, each with the same interface field definition. The interface field represents the values and procedures outlined above. The rule modules are stored in libraries for execution as needed. The fact that the data section of a module may be assigned to a pointer variable gives us the ability to unite the control store and the data store.

There are two possible methods of evaluating the rules. The first is to create instances of all the rules, and save the pointers to the correct MODULE as part of the data structure. The second is to create an instance of a rule only when it is necessary to evaluate that particular rule, and to dispose of the rule when it is no longer needed. The first method is much faster than the second, since at the time of the consultation the code for all rules is in core. The second method is extremely core efficient, however, since only a few rules will be active at any given time. This method is particularly suited for a small machine environment.

One of the advantages of this MAINSAIL approach is that there is no linking step in compiling modules as in other algorithmic languages (e.g., SAIL). Thus rule modules may be changed at will without relinking the entire system. The cost for the flexibility MAINSAIL offers is the overhead of intermodule calls in a dynamic memory environment - modules do not always load at the same address.

SAIL/LEAP Implementation

The LEAP package of SAIL offers yet another approach (4). Among the LEAP facilities are procedures which "assign" an ITEM (the basic element of LEAP) with a SAIL procedure and "apply" executes the procedure associated with an ITEM. Since ITEM's are part of the data store, and can be manipulated in data structures, this allows the needed interactions of control store and data store.

The strategy in SAIL is to associate an ITEM with each of the basic procedures for each rule. These ITEM's are stored so they can be retrieved when the appropriate rule is invoked. There are several points where special care must be taken. The program to modify the knowledge base must update several files in this implementation; 1) a header file which contains the declaration of all ITEM's used in the system, 2) an initialization procedure which creates all the triples, performs all the assigns, etc. needed, and 3) the actual code for the rule.

The advantages of the SAIL version is a somewhat more direct mapping of the data to a single procedure representing a rule. Adding a rule will be more cumbersome. The new rule can be stored in a file by itself, and required in the initialization module as a REQUIRE'd LOAD!MODULE. Thus only the new rule and the initialization need to be recompiled but the whole system must be relinked.

---

(4) LEAP is a facility added to SAIL for the associative storage, retrieval, and manipulation of objects. See Feldman, J.A. and Rovner, P.D., "An ALGOL-Based Associative Language," CACM 12, 8, August 1969.

Summary of Preliminary Results

A highly stripped down, seven rule "MYCIN" system has been successfully emulated using each of these approaches. These systems merely indicated the workability of each approach but the system was not complex enough to draw any quantitative conclusions about relative efficiency. Each of the methods has its advantages and disadvantages. SAIL, like LISP, requires a large core image and currently runs only on PDP-10 systems. The MAINSAIL version seems well suited to a small machine environment through the high modularization and dynamic memory management. By use of a virtual data structure, the rule interpretation approach could be made to run in a smaller core image as well.

It must be noted that no implementation of MYCIN in an algorithmic language will maintain the full flexibility of the INTERLISP version for rule changes, control structure experimentation, and debugging. One must be willing to trade the flexibility of a development system requiring large resources for a more fixed production oriented system with substantially smaller demands.

2.1.2.6 USER SOFTWARE AND INTRA-COMMUNITY COMMUNICATION

We have continued to assemble and maintain a broad range of utilities and user support software. These include operational aids, statistics packages, DEC-supplied programs, improvements to the TOPS-10 emulator, text editors, text search programs, file space management programs, graphics support, a batch program execution monitor, text formatting and justification assistance, and magnetic tape conversion aids. Over the past year we have made changes and updates to more than 60 programs in this stable. While many of these changes were maintenance bug fixes, major improvements were made to SPELL, MACRO, BACKUP, DIABLO, tape service programs, VIEW, and the user spooler interface. In addition we have brought up a number of new programs including PASCAL (DECUS), EMACS (a display oriented editor from MIT - installed by McMahon at SRI), overload control information programs, MACLISP (MIT), and FAIL. Changes are in progress to the bulletin board system to allow string searches in the "subject" and "body" of bulletins to find information of interest and to allow general wild card specifications within strings.

2.1.2.7 DOCUMENTATION AND EDUCATION

We have spent considerable effort to develop, maintain, and facilitate access to our documentation so as to accurately reflect available software. The HELP and Bulletin Board systems have been important in this effort. As subsystems are updated, we generally publish a bulletin or small document describing the changes. As more and more changes occur, it becomes harder and harder for users to track down all of the change pointers. We are in the process of reviewing the existing documentation system again for compatibility with the programs now on line and to integrate changes into the main documents. This will also be done with a view toward developing better tools for maintaining up-to-date documentation.

2.1.2.8 SOFTWARE COMPATIBILITY AND SHARING

At SUMEX-AIM we firmly believe in importing rather than reinventing software where possible. As noted above, a number of the packages we have brought up are from outside groups. Many avenues exist for sharing between the system staff, various user projects, other facilities, and vendors. The advent of fast and convenient communication facilities coupling communities of computer facilities has made possible effective intergroup cooperation and decentralized maintenance of software packages. The TENEX sites on the ARPANET have been a good model for this kind of exchange based on a functional division of labor and expertise. The other major advantage is that as a by-product of the constant communication about particular software, personal connections between staff members of the various sites develop. These connections serve to pass general information about software tools and to encourage the exchange of ideas among the sites. Certain common problems are now regularly discussed on a multi-site level. We continue to draw significant amounts of system software from other ARPANET sites, reciprocating with our own local developments. Interactions have included mutual backup support, hardware configuration experience, operating system enhancements, utility or language software, and user project collaborations. We have been able to import many new pieces of software and improvements to existing ones in this way. Examples of imported software include the message manipulation program MSG, TENEX SAIL, TENEX SOS, INTERLISP, the RECORD program, ARPANET host tables, and many others. Reciprocally, we have exported our contributions such as the drum page migration system, KI-10 page table efficiency improvements, GTJFN enhancements, PUB macro files, the bulletin board system, MAINSAIL, SPELL, SNDMSG enhancements, our BATCH monitor, and improved SA-10 software.

2.1.3 RESOURCE MANAGEMENT

2.1.3.1 ORGANIZATION

The SUMEX-AIM resource is administered within the Genetics Department of the Stanford University Medical School. Its mission, locally and nationally, entails both the recruitment of appropriate research projects interested in medical AI applications and the catalysis of interactions among these groups and the broader medical community. User projects are separately funded and autonomous in their management. They are selected for access to SUMEX on the basis of their scientific and medical merits as well as their commitment to the community goals of SUMEX. Currently active projects span a broad range of application areas such as clinical diagnostic consultation, molecular biochemistry, belief systems modeling, mental function modeling, and instrument data interpretation (descriptions of the individual collaborative projects are in Section 4 beginning on page 61).

Early this year it was announced that Professor Lederberg had been named president of Rockefeller University. Whereas the SUMEX staff at Stanford will miss the face-to-face contacts of his involvement in SUMEX-AIM, his relocation may even broaden and strengthen the biomedical research base that will be represented in our AI applications. Professor Lederberg has expressed a strong, continuing commitment to medical AI applications and to SUMEX. The network and message facilities provide a mechanism to continue his close participation in this research and AIM Executive Committee activities.

The depth of the Stanford multi-disciplinary support of SUMEX-AIM has been a key asset in being able to bridge this management transition. Professor Edward Feigenbaum, who is chairman of the Stanford Computer Science Department and has long been the co-principal investigator of SUMEX-AIM, will take over as PI. Professor Stanley Cohen, who has been PI of the MYCIN project and on the Stanford SUMEX advisory committee, will provide the biomedical ties and coordination with the Stanford Medical School and projects. The new management team is committed to sustaining the active development of the SUMEX-AIM resource and community.

2.1.3.2 MANAGEMENT COMMITTEES

As the SUMEX-AIM project is a multilateral undertaking by its very nature, we have created several management committees to assist in administering the various portions of the SUMEX resource. As defined in the SUMEX-AIM management plan adopted at the time the initial resource grant was awarded, the available facility capacity is allocated 40% to Stanford Medical School projects, 40% to national projects, and 20% to common system development and related functions. Within the Stanford aliquot, Dr. Lederberg and BRP have established an advisory committee to assist in selecting and allocating resources among projects appropriate to the SUMEX mission. The current membership of this committee is listed in Appendix IV.

For the national community, two committees serve complementary functions. An Executive Committee oversees the operations of the resource as related to national users and makes the final decisions on authorizing admission for projects. It also establishes policies for resource allocation and approves plans for resource development and augmentation within the national portion of SUMEX (e.g., hardware upgrades, MAINSAIL development priorities, etc.). The Executive Committee oversees the planning and implementation of the AIM Workshop series currently implemented under Prof. S. Amarel of Rutgers University and assures coordination with other AIM activities as well. The committee will play a key role in assessing the possible need for additional future AIM community computing resources and in deciding the optimal placement and management of such facilities. The current membership of the Executive committee is listed in Appendix IV.

Reporting to the Executive Committee, an Advisory Group represents the interests of medical and computer science research relevant to AIM goals. The Advisory Group serves several functions in advising the Executive Committee; 1) recruiting appropriate medical/computer science projects, 2) reviewing and recommending priorities for allocation of resource capacity to specific projects based on scientific quality and medical relevance, and 3) recommending policies and development goals for the resource. The current Advisory Group membership is given in Appendix IV.

These committees have actively functioned in support of the resource. Except for the meetings held during the AIM workshops, the committees have "met" by messages, net-mail, and telephone conference owing to the size of the groups and to save the time and expense of personal travel to meet face to face. The telephone meetings, in conjunction with terminal access to related text materials, have served quite well in accomplishing the agenda business and facilitate greatly the arrangement of meetings. Other solicitations of advice requiring review of sizable written proposals are done by mail.

We will continue to work with the management committees to recruit the additional high quality projects which can be accommodated and to evolve resource allocation policies which appropriately reflect assigned priorities and project needs. We hope to make more generally available information about the various projects both inside and outside of the community and thereby to promote the kinds of exchanges exemplified earlier and made possible by network facilities.

2.1.3.3 NEW PROJECT RECRUITING

The SUMEX-AIM resource has been announced through a variety of media as well as by correspondence, contacts of NIH-BRP with a variety of prospective grantees who use computers, and contacts by our own staff and committee members. The number of formal projects that have been admitted to SUMEX has more than doubled since the start of the project; others are working tentatively as pilot projects or are under review.

We have prepared a variety of materials for the new user ranging from general information such as is contained in a SUMEX-AIM overview brochure to more detailed information and guidelines for determining whether a user project is appropriate for the SUMEX-AIM resource. Dr. E. Levinthal has prepared a questionnaire to assist users seriously considering applying for access to SUMEX-AIM. Pilot project categories have been established both within the Stanford and national aliquots of the facility capacity to assist and encourage projects just formulating possible AIM proposals pending their application for funding support and in parallel formal application for access to SUMEX. Pilot projects are approved for access for limited periods of time after preliminary review by the Stanford or AIM Advisory Group as appropriate to the origin of the project.

These contacts have sometimes done much more than provide support for already formulated programs. For example, Prof. Feigenbaum's group at Stanford has initiated a major collaborative effort with Dr. Osborn's group at the Institutes of Medical Sciences in San Francisco. This project in "Pulmonary Function Monitoring and Ventilator Management - PUFF/VM" (see Section 4.1.6 on page 93) originated as a pilot request to use MLAB in a small way for modeling. Subsequently the AI potentialities of this domain were recognized by Feigenbaum, Nii, and Osborn who have submitted a joint proposal to NIH and have a pilot status at present. This summer Dr. John Kunz from Dr. Osborn's laboratory is planning to spend half time at Stanford to learn more about AI research and to participate more closely in the development of the PUFF/VM program.

The following lists the fully authorized projects currently comprising the SUMEX-AIM community (see Section 4 for more detailed descriptions). The nucleus of five projects that were authorized at the initial funding of the resource in December 1973 are marked by "<*>".

<u>National Community</u> -

1) Acquisition of Cognitive Procedures (ACT); Dr. J. Anderson (Yale University)

2) Chemical Synthesis Project (SECS); Dr. T. Wipke (University of California at Santa Cruz)

<*> 3) Higher Mental Functions Project; K. Colby, M.D. (University of California at Los Angeles)

4) INTERNIST Project; J. Myers, M.D. and Dr. H. Pople (University of Pittsburgh)

5) Medical Information Systems Laboratory (MISL); J. Wilensky, M.D. and Dr. B. McCormick (University of Illinois at Chicago Circle)

6) Pulmonary Function Project (PUFF/VM); J. Osborn, M.D. (Institutes of Medical Sciences, San Francisco) and Dr. E. Feigenbaum (Stanford University)

<*> 7) Rutgers Computers in Biomedicine; Dr. S. Amarel (Rutgers University)

8) Simulation of Comprehension Processes; Drs. J. Greeno and A. Lesgold (University of Pittsburgh)

<u>Stanford Community</u> -

1) AI Handbook Project; Dr. E. Feigenbaum

<*> 2) DENDRAL Project; Drs. C. Djerassi, J. Lederberg, and E. Feigenbaum

3) Generalization of AI Tools (AGE); Dr. E. Feigenbaum

4) Large Multi-processor Arrays (HYDROID); Dr. G. Wiederhold

5) Molecular Genetics Project (MOLGEN); Drs. J. Lederberg and E. Feigenbaum (Stanford) and N. Martin (University of New Mexico)

<*> 6) MYCIN Project; S. Cohen, M.D. and Dr. B. Buchanan

<*> 7) Protein Structure Modelling; Drs. E. Feigenbaum and R. Engelmore

As an additional aid to new projects or collaborators with existing projects, we provide a limited amount of funds for use to support terminals and communications needs of users without access to such equipment. We are currently leasing 6 terminals and 4 modems for users as well as 4 foreign exchange lines to better couple the Rutgers project into the TYMNET and a leased line between Stanford and U. C. Santa Cruz for the Chemical Synthesis project.

2.1.3.4 <u>STANFORD COMMUNITY BUILDING</u>

The Stanford community has undertaken several internal efforts to encourage interactions and sharing between the projects centered here. Professor Feigenbaum organized a project with the goal of assembling a handbook of AI concepts, techniques, and current state-of-the-art. This project has had enthusiastic support from the students and substantial progress made in preparing many sections of the handbook (see Section 4.2.1 on page 123 for more details).

Weekly informal lunch meetings (SIGLUNCH) are also held between community members to discuss general AI topics, concerns and progress of individual projects, or system problems as appropriate as well as having a number of outside invited speakers.

2.1.3.5 <u>AIM WORKSHOP SUPPORT</u>

The Rutgers Computers in Biomedicine resource (under Dr. Saul Amarel) has organized a series of workshops devoted to a range of topics related to artificial intelligence research, medical needs, and resource sharing policies within NIH. Meetings have been held for the past several years at Rutgers and another is planned for this summer. The SUMEX facility has acted as a prime computing base for the workshop demonstrations. We expect to continue this support for future workshops. The AIM workshops provide much useful information about the strengths and weaknesses of the performance programs both in terms of criticisms from other AI projects and in terms of the needs of practicing medical people. We plan to continue to use this experience to guide the community building aspects of SUMEX-AIM.

2.1.3.6 RESOURCE ALLOCATION POLICIES

As the SUMEX facility has become increasingly loaded, a number of diverse and conflicting demands have arisen which require controlled allocation of critical facility resources (file space and central processor time). We have already spelled out a policy for file space management; an allocation of file storage is defined for each authorized project in conjunction with the management committees. This allocation is divided among project members in any way desired by the individual principal investigators. System allocation enforcement is implemented by project each week. As the weekly file dump is done, if the aggregate space in use by a project is over its allocation, files are archived from user directories over allocation until the project is within its allocation.

We have recently implemented system scheduling controls to attempt to maintain the 40:40:20 balance in terms of CPU utilization (see page 14) and to avoid system and user inefficiencies during overload conditions. The initial complement of user projects justifying the SUMEX resource was centered to a large extent at Stanford. Over the past five years of the SUMEX grant, a substantial growth in the number of national projects was realized. During the same time the Stanford group of projects has matured as well and in practice the 40:40 split between Stanford and non-Stanford projects is not ideally realized (see Figure 11 on page 47 and the tables of recent project usage on page 50). Our job scheduling controls bias the allocation of CPU time based on percent time consumed relative to the time allocated over the 40:40:20 community split. The controls are "soft" however in that they do not waste computer cycles if users below their allocated percentages are not on the system to consume the cycles. The operating disparity in CPU use to date reflects a substantial difference in demand between the Stanford community and the developing national projects, rather than inequity of access. For example, the Stanford utilization is spread over a large part of the 24-hour cycle, while national-AIM users tend to be more sensitive to local prime-time constraints. (The 3-hour time zone phase shift across the continent is of substantial help in load balancing.) During peak times under the new overload controls, the Stanford community still experiences mutual contentions and delays while the AIM group has relatively open access to the system. For the present, we propose to continue our policy of "soft" allocation enforcement for the fair split of resource capacity.

Our system also categorizes users in terms of access privileges. These comprise fully authorized users, pilot projects, guests, and network visitors in descending order of system capabilities. We want to encourage bona fide medical and health research people to experiment with the various programs available with a minimum of red tape while not allowing unauthenticated users to bypass the advisory group screening procedures by coming on as guests. So far we have had relatively little abuse compared to what other network sites have experienced, perhaps on account of the personal attention that senior staff gives to the logon records, and to other security measures. However, the experience of most other computer managers behooves us to be cautious about being as wide open as might be preferred for informal service to pilot efforts and demonstrations. We will continue developing this mechanism in conjunction with management committee policy decisions.

2.1.4 FUTURE PLANS

This next year will be the first of the 3 year renewal grant term. The principal goals of our work are outlined below. Objectives for the individual collaborating projects are discussed in their respective reports (see Section 4 on page 61).

1) RESOURCE OPERATIONS

We will continue to make available to the SUMEX-AIM communities an effective, state-of-the-art facility to support the development of medical AI programs and to facilitate collaborations between community members. Goals include:

a) Assure a smooth transition in project management as Professor Lederberg moves to Rockefeller University.

b) Continue development of the existing KI-TENEX facility to maximize effectiveness for community use. We expect to continue improving system efficiency, allocation controls, subsystem software, documentation facilities, and communications facilities. We will complete the evaluation of the TELENET network as a more cost-effective source of communication services. Another key issue will be how best to maintain software compatibility between TENEX and the newer releases of DEC's TOPS-20. This may entail another "compatibility package" to translate system calls from one system to the other.

c) Recruit new applications and projects to broaden the range of high quality medical AI applications. We look forward to Prof. Lederberg's efforts at Rockefeller University to try to stimulate new projects as well as others that might be suggested by advisory group members or other contacts.

d) We plan to work closely with other AIM resource nodes, such as the one being implemented at Rutgers this summer, to ensure effective community support between the facilities and to take advantage of expertise in various user groups for system and user software development.

e) We plan to finish the preliminary evaluation of the new DEC 2020 system and to make a recommendation for acquiring one by the end of calendar 1978. The council-approved budget allocation for this machine in year 07. We expect the technological rationale and community need for the use of such a machine for increased capacity and an effective software export mode to mature in year 06. Thus it would be desirable to move this expenditure forward. This may not be possible within NIH appropriation limitations and so we would have to defer delivery until year 07. As part of the acquisition of the 2020 system, we expect to investigate the many issues that will arise from the decentralization such machines will bring. The availability of these machines bodes many advantages for effective support of community computing needs but dangers as well of decreased sharing and software compatibility.

2) TRAINING AND EDUCATION

Within our resources, we will continue to assist new and established user projects in gaining access to SUMEX-AIM facilities. Collaborating projects will provide their own manpower and expertise for the development and dissemination of their AI programs.

a) We will continue to provide a high standard of system documentation and limited staff assistance for user problems.

b) Council disapproved our plan to support a "visiting scientist" position to bring selected investigators in contact with on-going AI projects. Funds were approved to support "collaborative linkages". These will continue to be used to facilitate project communications with the resource.

c) We will provide continued support for the AIM workshop activities in the form of demonstration support, participation in workshop discussions, and assistance for potential pilot users in understanding the SUMEX-AIM community.

3) CORE RESEARCH

Our core research efforts for the next year will emphasize the research work discussed in our proposal but the level of effort will reflect the budget cuts recommended by Council. This effect will be particularly hard felt in the MAINSAIL project.

a) We will provide core research support to about 1.6 staff FTE's and 1 graduate research assistant for the documentation and generalization of AI tools developed in the context of particular applications projects. This work will complement the on-going project developments by providing a link to make results available to the entire community. We plan to partially support the AGE project and the AI handbook project. The detailed research goals of these projects are summarized in Section 4.2.1 and Section 4.2.3.

b) Within the council-approved manpower level for MAINSAIL (2 FTE's), we will only be able to complete a demonstration of the MAINSAIL system. A wide distribution of the language, credible support of a user community, and investigation of implementations for other target machines are well beyond this level of effort. For the next year we plan to complete a debugged compiler that will run on a 28K PDP-11, tutorial and reference manuals (including procedures for bringing up MAINSAIL on new target machines), an interactive symbolic debugger (providing breakpoints, variable examination, single stepping, etc.), and documentation of the key design issues encountered in defining a machine-independent language. Also over this year we will investigate ways in which this demonstration version of MAINSAIL could be transferred to an environment with the necessary resources to extend, distribute, and maintain it properly.

2.2 SUMMARY OF RESOURCE USAGE

The following data give an overview of SUMEX-AIM resource usage. There are five subsections containing data respectively for 1) system loading, 2) system efficiency, 3) resource use by community, 4) resource use by project, and 5) network use.

2.2.1 SYSTEM LOADING

The following plots display several different aspects of system loading over life of the project. These include total CPU time delivered per month, the "peak" number of jobs logged in, and the "peak" load average. The term "peak" refers to the peak of the monthly diurnal loading curve for each variable which in turn is the average of the individual daily diurnal curves. Thus, "peak" values are quite representative of average monthly peak loading and do not reflect individual days. These data show well the continued growth of SUMEX use and the self-limiting saturation effect of system load average. Since late 1976, when the dual processor capacity became fully used, the peak daily load average has remained at about 6. This is a measure of the user capacity of our current hardware configuration and the mix of AI programs.

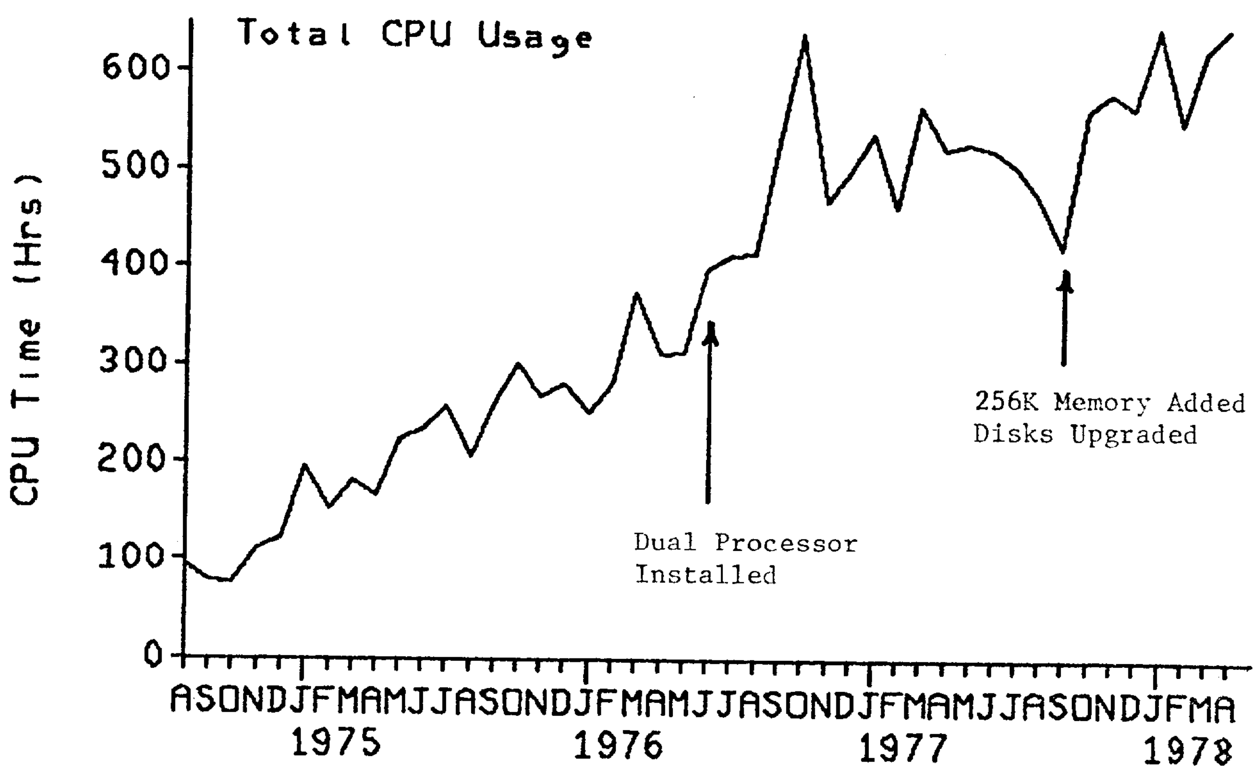

Figure 6. CPU Time Consumed by Month

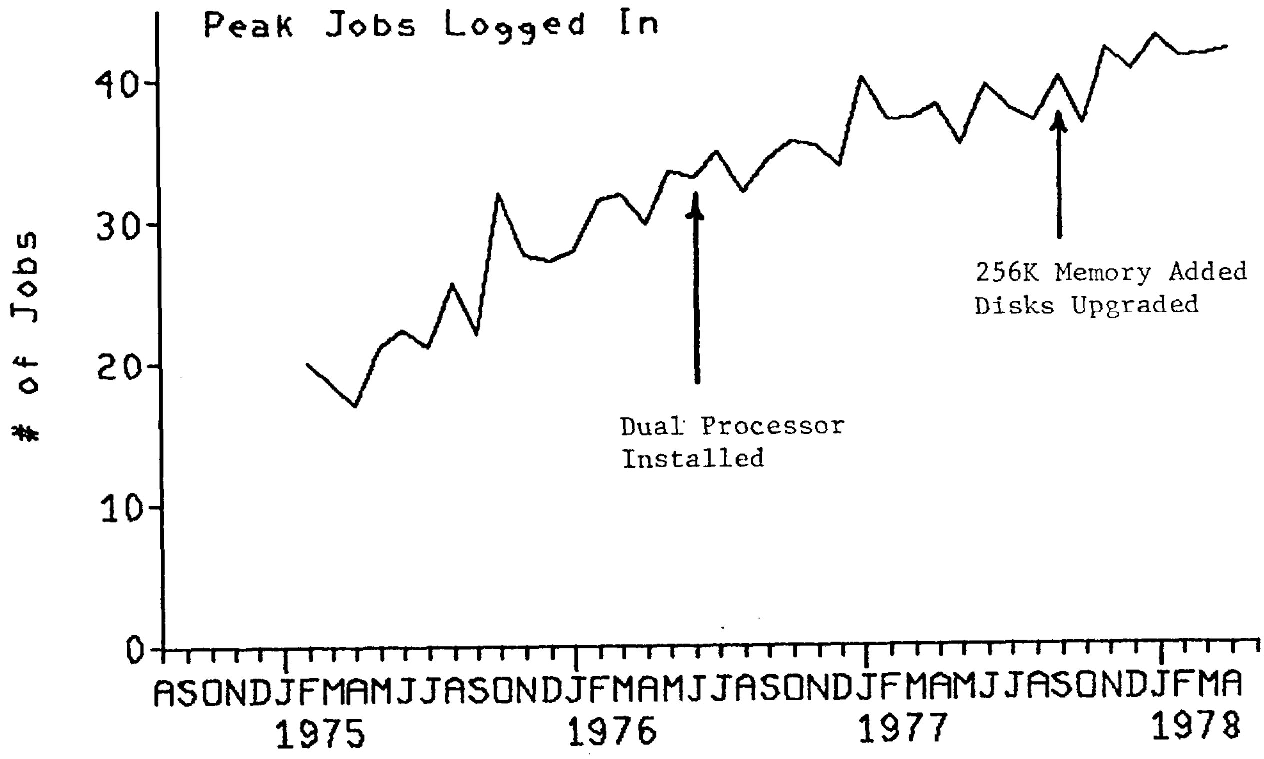

Figure 7. Peak Number of Jobs by Month

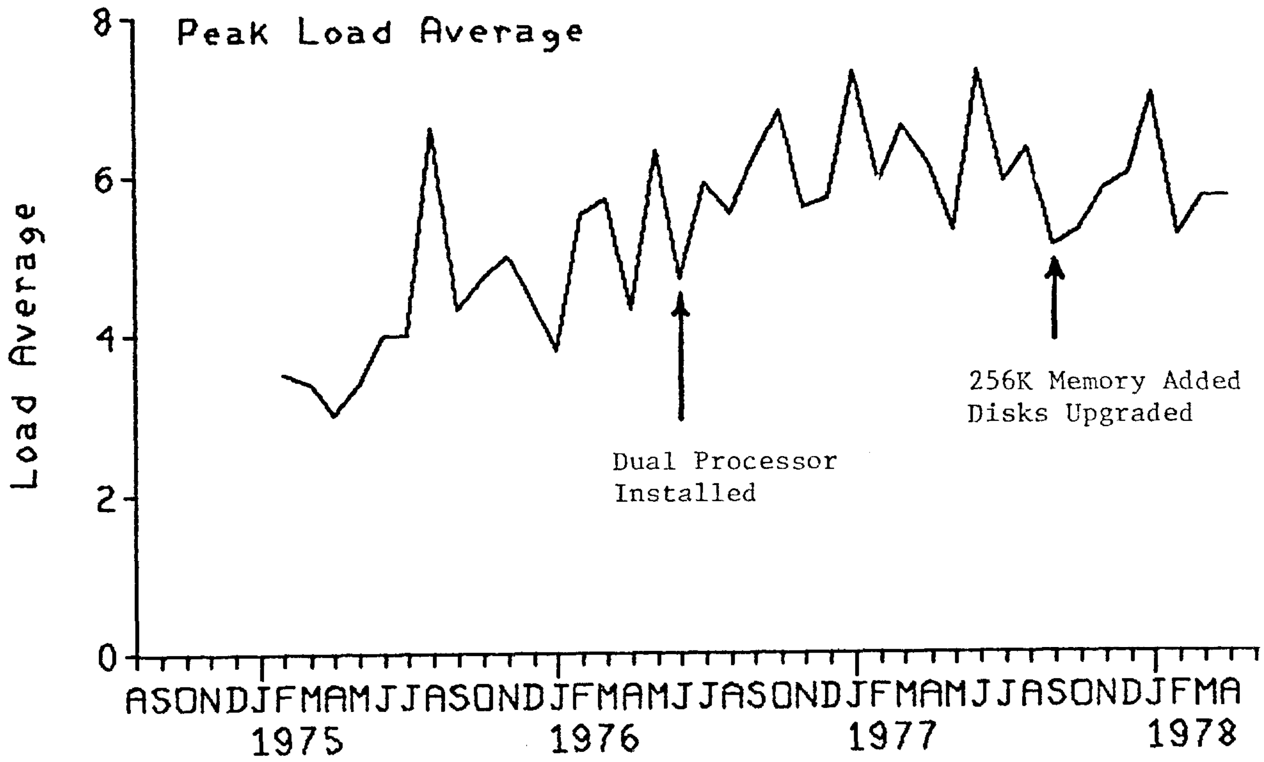

Figure 8. Peak Load Average by Month

2.2.2 SYSTEM EFFICIENCY

The following plots show two measures of system overhead and the influence of hardware augmentations on them. The first is "total overhead" which includes scheduler time, I/O wait, and core management time. The second shows "page trap" time which is charged to user run time but reflects lost time in working set management for each job. Note the sharp rise in overhead with the introduction of the dual processor caused by the increased memory contention. This overhead drops back to single processor levels after we doubled memory. The peak around February/March 1975 is anomalous and reflects testing of the drum and disk systems during installation.

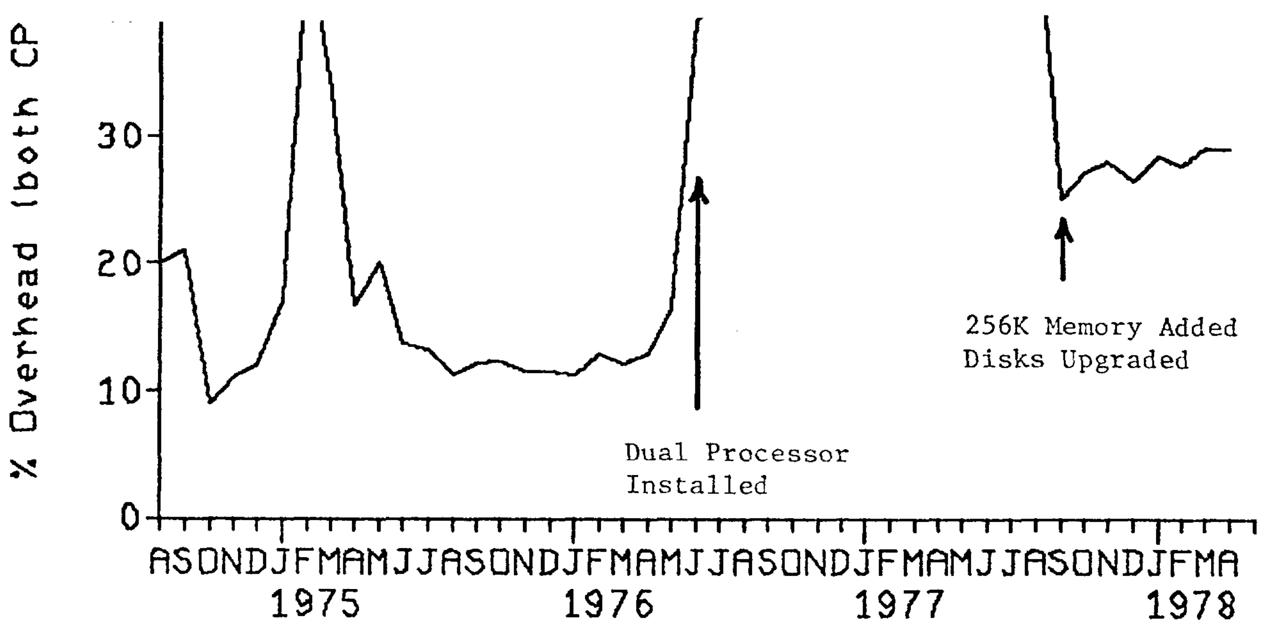

Figure 9. System Overhead by Month

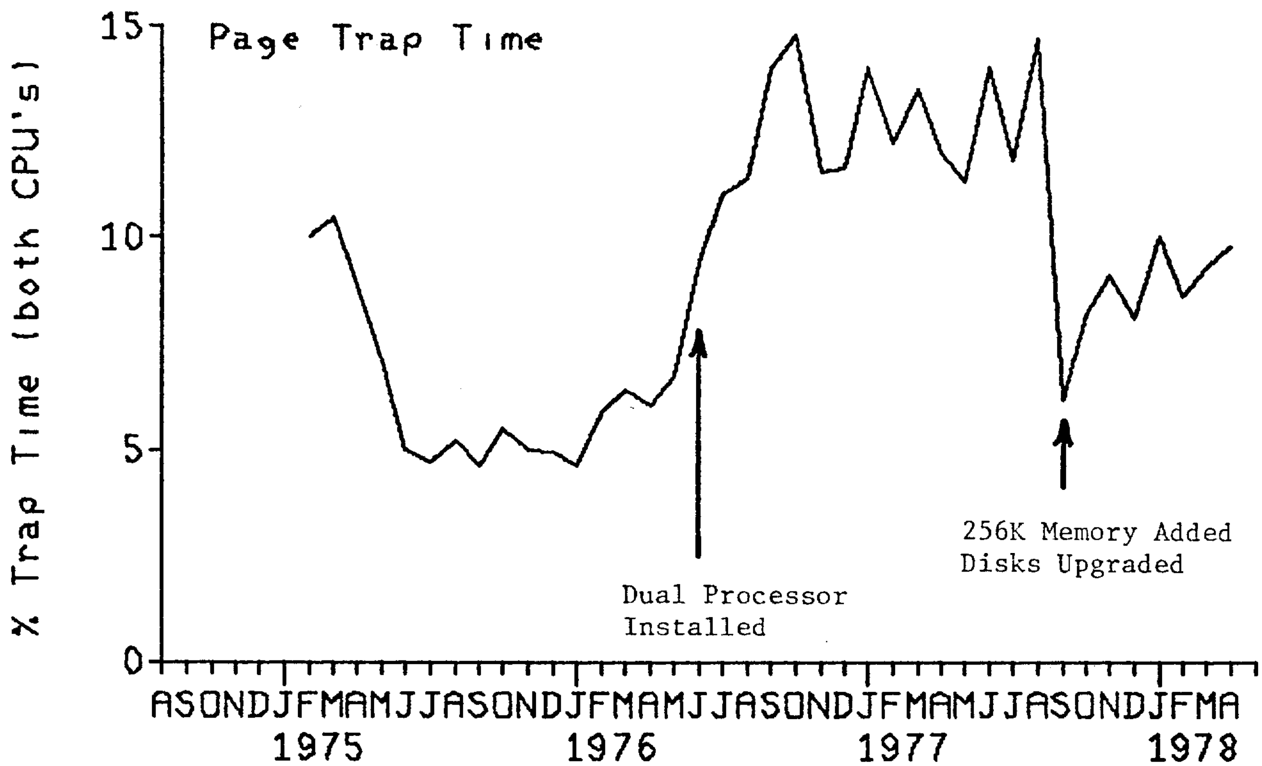

Figure 10. Page Trap Time by Month

2.2.3 RELATIVE SYSTEM LOADING BY COMMUNITY

The SUMEX resource is divided, for administrative purposes, into 3 major communities: user projects based at the Stanford Medical School, user projects based outside of Stanford (national AIM projects), and common systems development efforts. As defined in the resource management plan approved by BRP at the start of the project, the available system CPU capacity and file space resources are divided between these communities as follows:

Stanford	40%
AIM	40%
Staff	20%

The "available" resources to be divided up in this way are those remaining after various monitor and community-wide functions are accounted for. These include such things as job scheduling, overhead, network service, file space for subsystems, documentation, etc.

The monthly usage of CPU and file space resources for each of these three communities relative to their respective aliquots is shown in the plots in Figure 11 and Figure 12. Terminal connect time is shown in Figure 13. It is clear that the Stanford projects have held an edge in system usage despite our efforts at resource allocation and the substantial voluntary efforts by the Stanford community to utilize non-prime hours. This reflects the development of the Stanford group of projects relative to those getting started on the national side and has correspondingly accounted for much of the progress in AI program development to date.

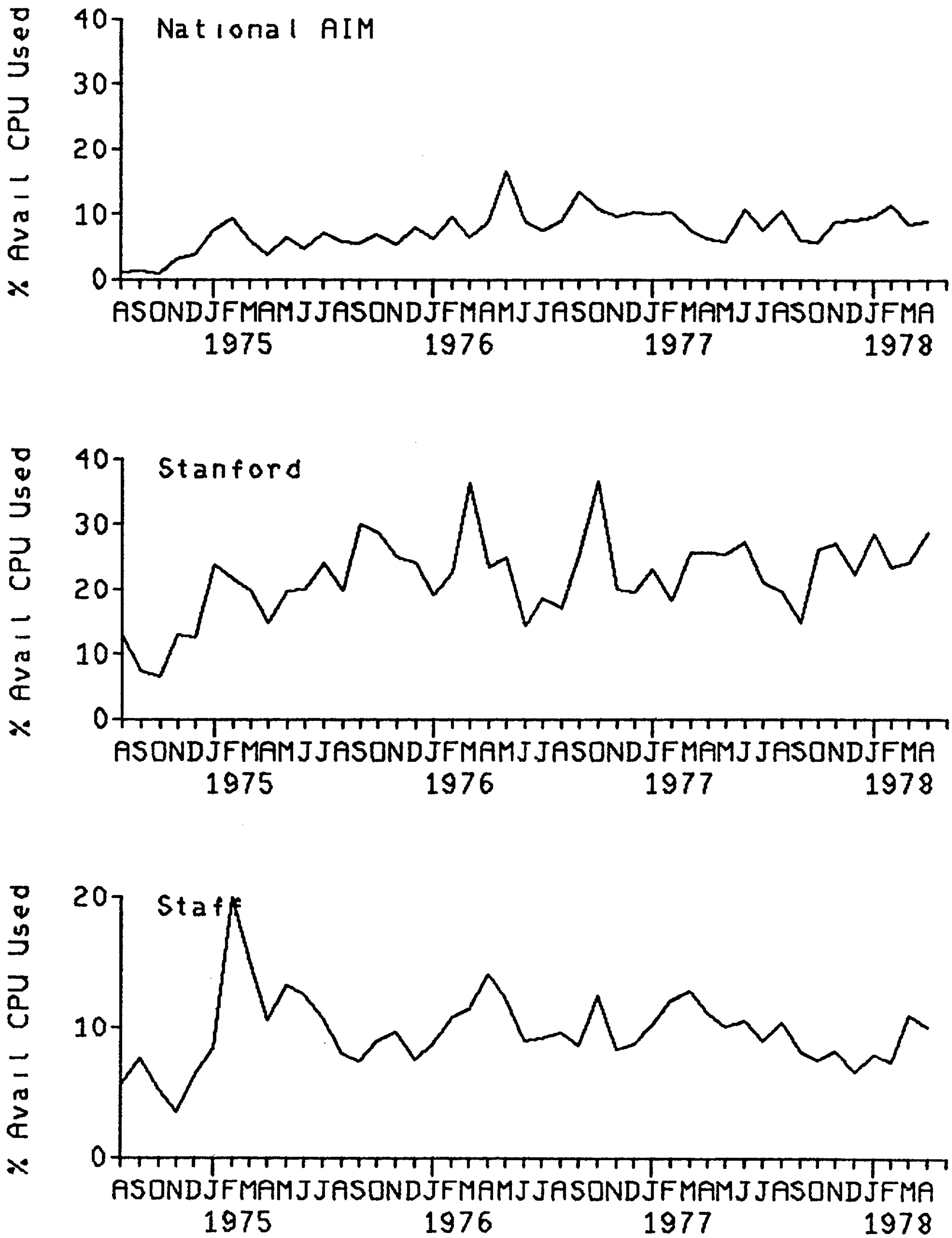

Figure 11. Monthly CPU Usage by Community

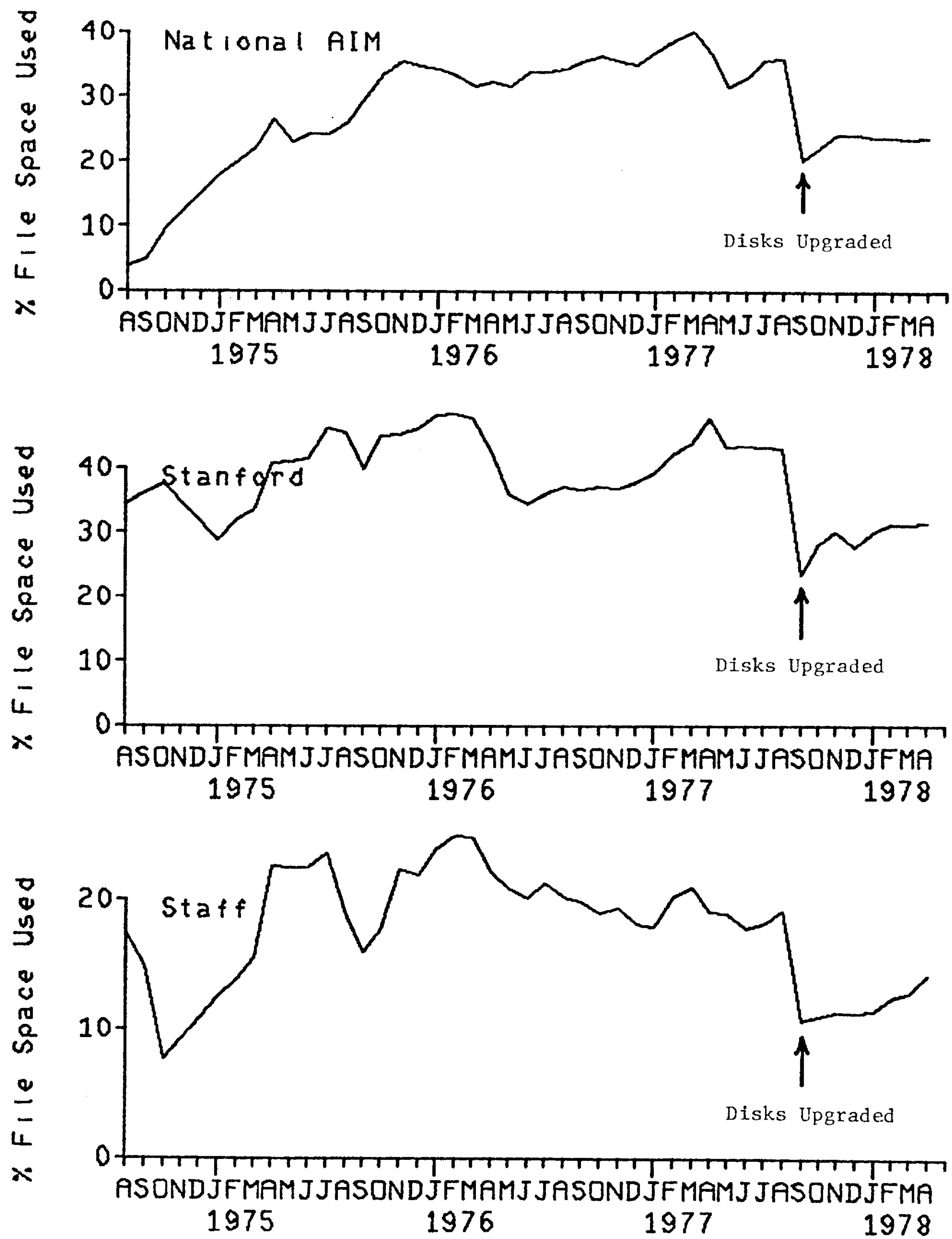

Figure 12. Monthly File Space Usage by Community

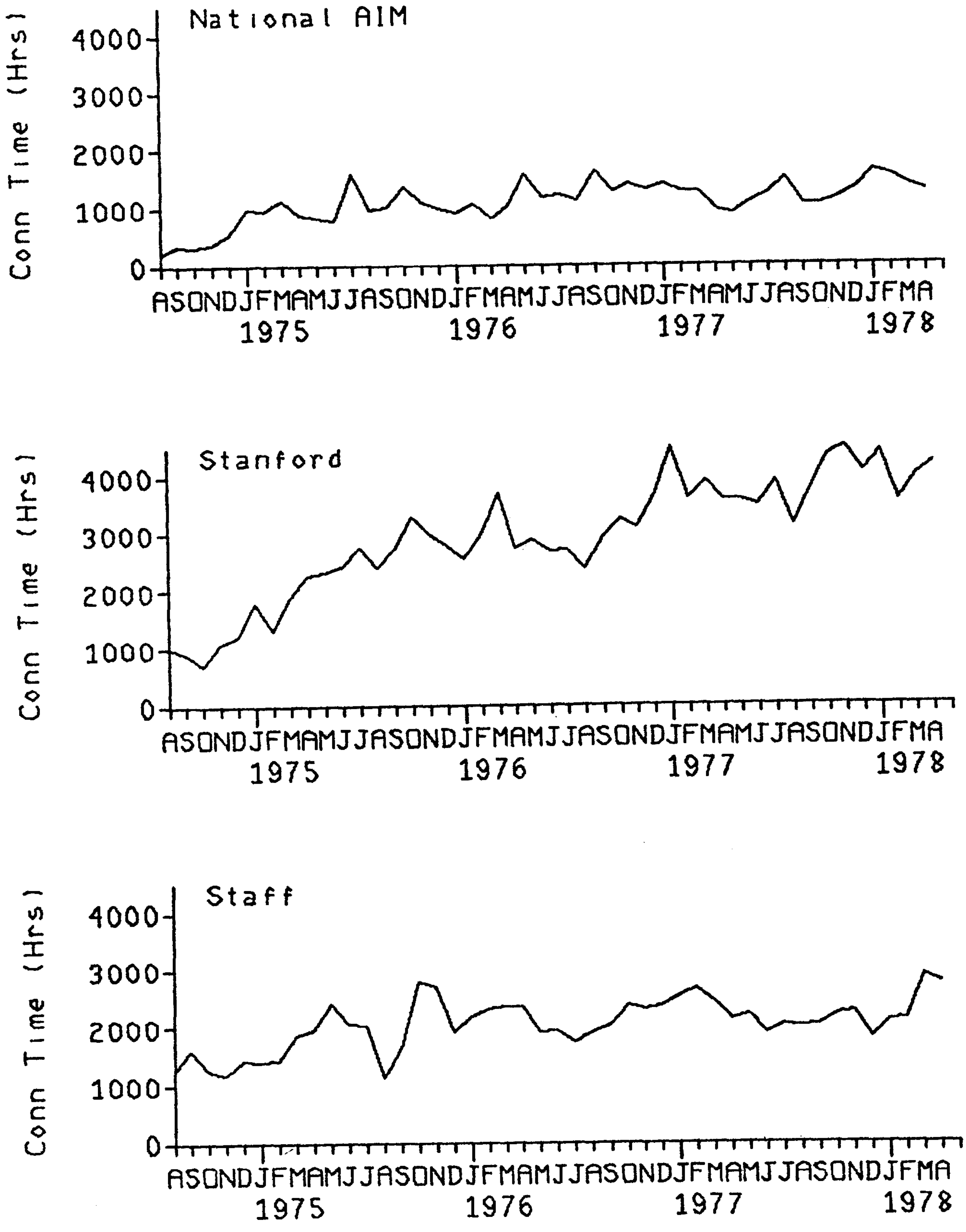

Figure 13. Monthly Terminal Connect Time by Community

2.2.4 INDIVIDUAL PROJECT AND COMMUNITY USAGE

The table following shows cumulative resource usage by project in the past grant year. The data displayed include a description of the operational funding sources (outside of SUMEX-supplied computing resources) for currently active projects, total CPU consumption by project (Hours), total terminal connect time by project (Hours), and average file space in use by project (Pages, 1 page = 512 computer words). These data were accumulated for each project for the months between May 1977 and April 1978. Again the well developed use of the resource by the Stanford community can be seen. It should be noted that the Stanford projects have voluntarily shifted a substantial part of their development work to non-prime time hours which is not shown in these cumulative data. It should also be noted that a significant part of the DENDRAL and MYCIN efforts, here charged to the Stanford aliquot, support development efforts dedicated to national community access to these systems. The actual demonstration and use of these programs by extramural users is charged to the national community in the "AIM USERS" category, however.

RESOURCE USE BY INDIVIDUAL PROJECT - 5/77 THROUGH 4/78

NATIONAL AIM COMMUNITY	CPU (Hours)	CONNECT (Hours)	FILE SPACE (Pages)
1) ACT PROJECT "Acquisition of Cognitive Procedures" John Anderson, Ph.D. Yale University ONR N0014-77-6-0242 (3.5 yrs. 3/77-9/80) 3/78-2/79 $90,000 (*)	153.83	2348.60	2016
2) SECS PROJECT "Simulation & Evaluation of Chemical Synthesis" W. Todd Wipke, Ph.D. U. California, Santa Cruz NIH RR-01059-01 (3 yrs. 7/77-6/80) 7/77-6/78 $94,602 NCI N01-CP-75816 (18 mos.) 3/77-9/78 $58,753 Bayer $5,000 E. Merck $1,500 Sandoz $2,500	360.14	4823.39	7602
3) HIGHER MENTAL FUNCTIONS "Computer Models in Psychiatry and Psychother." Kenneth Colby, M.D. UCLA Biobehavioral Sciences Program funding Proposals pending	48.21	724.20	2374
4) INTERNIST PROJECT "DIALOG: Computer Model of Diagnostic Logic" Jack Myers, M.D. Harry Pople, Ph.D. University of Pittsburgh BHRD MB-00144-04 (4 yrs. 7/74-6/78) 7/77-6/78 $101,000 NIH RR-01101-01 (3 yrs. 7/77-6/80) 7/77-6/78 $160,000	189.68	2735.83	6138

5)	MISL PROJECT "Medical Information Systems Laboratory" Morton Goldberg, M.D. Bruce McCormick, Ph.D. U. Illinois, Chicago Cir. US-PHS-MB00114-04 (4 yrs. 7/74-6/78) 7/77-6/78 $222,487	7.21	251.59	1001
6)	PUFF-VM PROJ (since 10/77) "Pulmonary Function Diag. & Ventilator Management" Edward Feigenbaum, Ph.D. Stanford University John Osborn, M.D. Inst. Medical Sciences, San Francisco NIH approved but unfunded	55.89	1395.85	1234
7)	RUTGERS PROJECT "Computers in Biomedicine" Saul Amarel, Ph.D. NIH RR-00643 (3 yrs. 12/77-11/80) 12/77-11/78 $505,823	32.87	500.10	8437
8)	SCP PROJECT (since 2/78) "Simulation of Comprehension Processes" James Greeno, Ph.D. Alan Lesgold, Ph.D. University of Pittsburgh ONR N0014-78-C-0022 (3 yrs. 10/77-9/80) 10/77-9/78 $62,616 OB-NIE-78-0115 12/77-11/78 $125,900	3.46	61.51	80
9)	AIM PILOT PROJECTS	85.63	1738.97	1911
10)	AIM Administration	16.30	449.84	1155
11)	AIM Users of DENDRAL and MYCIN	27.38	520.42	575
		------	--------	-----
	COMMUNITY TOTALS	980.60	15550.30	32523

STANFORD COMMUNITY	CPU (Hours)	CONNECT (Hours)	FILE SPACE (Pages)
1) AI HANDBOOK PROJECT Edward Feigenbaum, Ph.D. ARPA MDA-903-77-C-0322 (**) (partial support)	100.23	1667.92	1311
2) DENDRAL PROJECT "Resource Related Research Computers and Chemistry" Carl Djerassi, Ph.D. NIH RR-00612 (3 yrs. 5/77-4/80) 5/77-4/78 $218,580	1203.58	20060.60	17460
3) AGE PROJECT (since 9/77) "Generalization of AI Tools" Edward Feigenbaum, Ph.D. ARPA MDA-903-77-C-0322 (**) (partial support)	26.03	775.32	932
4) HYDROID PROJECT "Distributed Processing and Problem Solving" Gio Wiederhold, Ph.D. ARPA MDA-903-77-C-0322 (**)	117.36	2597.67	952
5) MOLGEN PROJECT "Experiment Planning System for Molecular Genetics" Edward Feigenbaum, Ph.D. Joshua Lederberg, Ph.D. NSF MCS76-11649 (2 yrs. 6/76-5/78) 6/77-5/78 $65,610 (*) Nancy Martin, Ph.D. U. New Mexico NSF MCS76-11935 (2 yrs. 7/76-6/78) Total award $68,000 (*)	197.17	3966.60	3125
6) MYCIN PROJECT "Computer-based Consult. in Clin. Therapeutics" Stanley N. Cohen, M.D. Bruce G. Buchanan, Ph.D. NSF MCS77-02712 (2 yrs. 6/77-5/79) 6/77-5/78 $32,357	487.35	6874.76	7892

7) PROTEIN STRUCT MODELING "Heuristic Comp. Applied to Prot. Crystallog." Edward Feigenbaum, Ph.D. NSF MCS-74-23461 (2 yrs. 5/77-4/79) Total award $150,200 (*)	174.21	2933.10	3499
8) PILOT PROJECTS	336.23	6885.61	3728
	-------	--------	-----
COMMUNITY TOTALS	2642.16	45761.58	38899

SUMEX STAFF	CPU (Hours)	CONNECT (Hours)	FILE SPACE (Pages)
1) Staff	661.53	20694.04	14317
2) MAINSAIL Development (since 9/77)	300.63	5125.66	2492
3) Staff affiliates, misc.	33.27	899.74	1219
	-------	---------	------
COMMUNITY TOTALS	995.43	26719.44	18028

SYSTEM OPERATIONS	CPU (Hours)	CONNECT (Hours)	FILE SPACE (Pages)
1) Operations	1986.55	78313.25	75657
	=======	=========	======
RESOURCE TOTALS	6604.74	166344.57	165107

* Award includes indirect costs. All other awards are reported as total direct costs only.

** Supported by a larger ARPA contract MDA-903-77-C-0322 awarded to the Stanford Computer Science Department for the Heuristic Programming Project for the period 8/77-9/79 at a funding level of $765,000 (incl. indirect costs).

2.2.5 NETWORK USAGE

These plots show total terminal connect time for TYMNET and ARPANET users by month since initial connection.

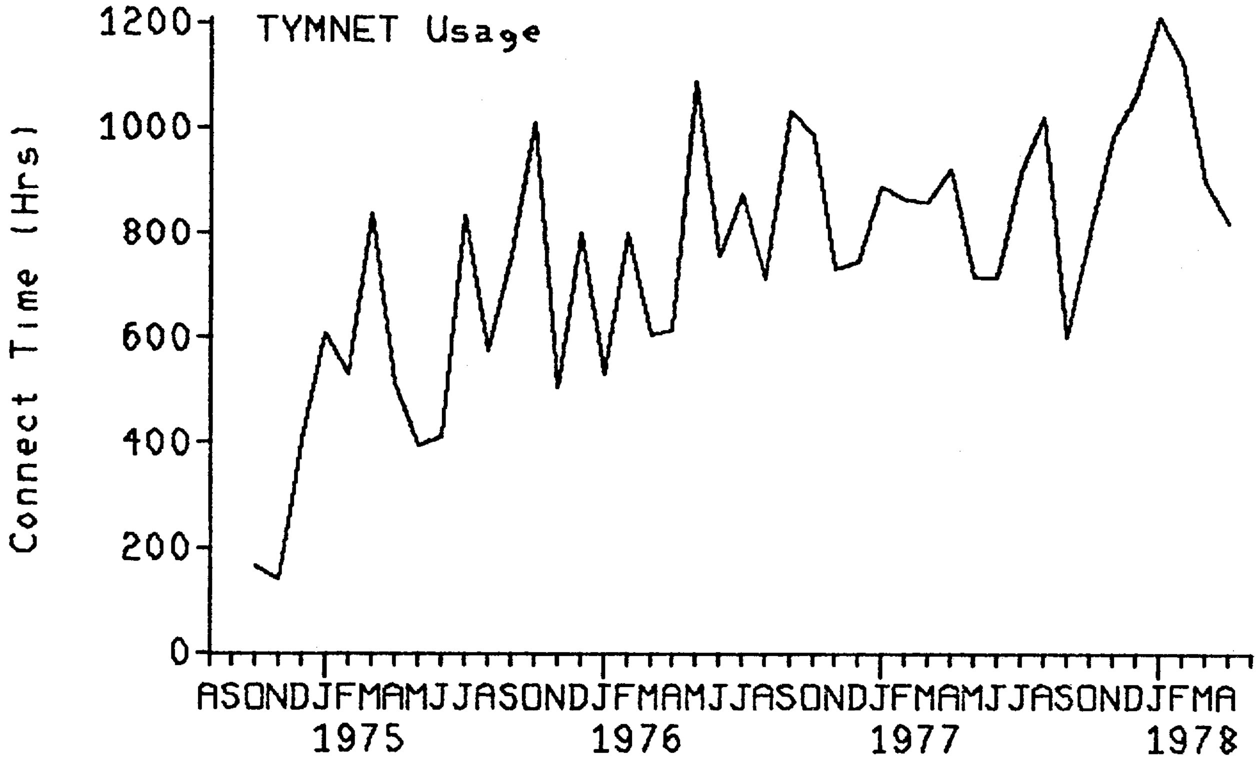

Figure 14. TYMNET Usage Data

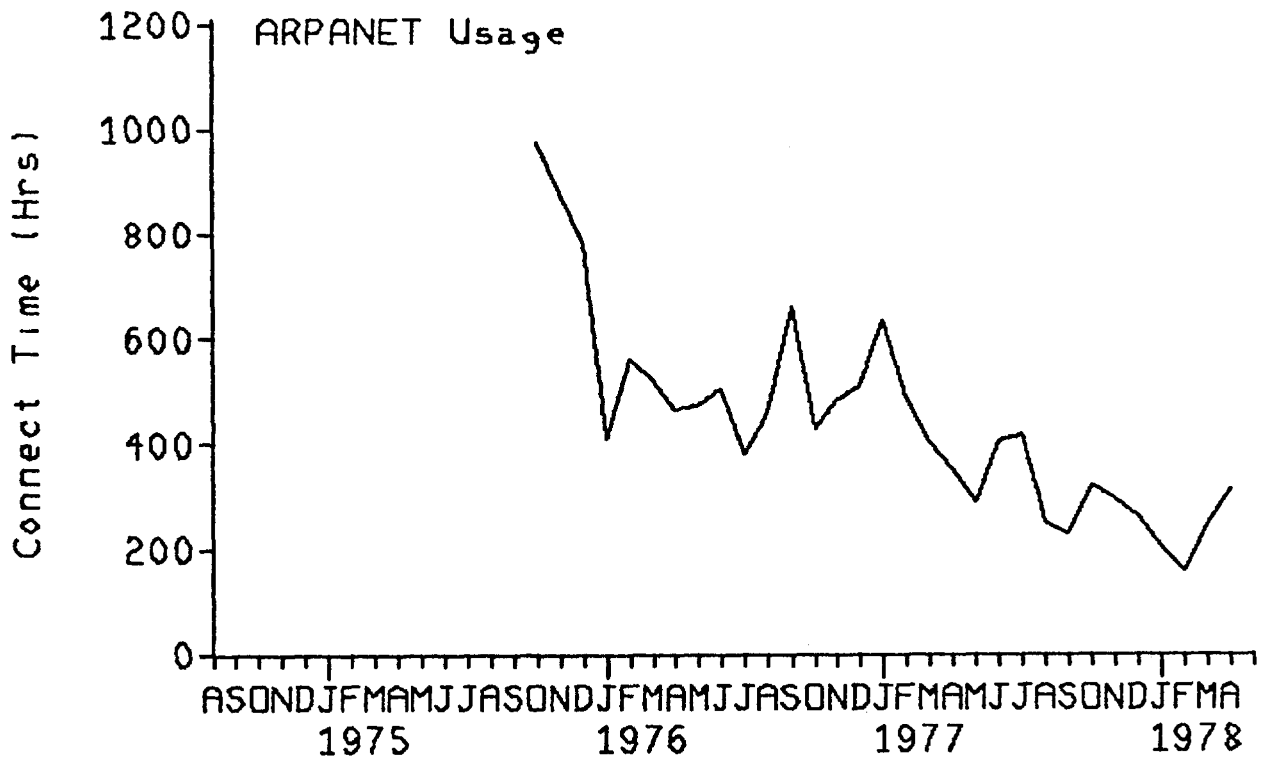

Figure 15. ARPANET Usage Data

2.3 RESOURCE EQUIPMENT SUMMARY

A complete inventory of resource equipment is attached separately as part of the budget material.

2.4 PUBLICATIONS

The following are publications for the SUMEX staff and have included papers describing the SUMEX-AIM resource and on-going research as well as documentation of system and program developments. Publications for individual collaborating projects are detailed in their respective reports (see Section 4 on page 61).

[1] Carhart, R.E., Johnson, S.M., Smith, D.H., Buchanan, B.G., Dromey, R.G., and Lederberg, J, "Networking and a Collaborative Research Community: a Case Study Using the DENDRAL Programs", ACS Symposium Series, Number 19, COMPUTER NETWORKING AND CHEMISTRY, Peter Lykos (Editor), 1975.

[2] Levinthal, E.C., Carhart, R.E., Johnson, S.M., and Lederberg, J., "When Computers Talk to Computers", Industrial Research, November 1975

[3] Wilcox, C. R., "MAINSAIL - A Machine-Independent Programming System," Proceedings of the DEC Users Society, Vol 2, No 4, Spring 1976.

[4] Wilcox, Clark R., "The MAINSAIL Project: Developing Tools for Software Portability," Proceedings, Computer Application in Medical Care, October, 1977, pp. 76-83.

[5] Lederberg, J. L., "Digital Communications and the Conduct of Science - THE NEW LITERACY," Accepted for publication, Proc. IEEE special issue on packet-switched communications.

Mr. Clark Wilcox also chaired the session on "Languages for Portability" at the DECUS DECsystem10 Spring '76 Symposium.

In addition as reported earlier, a substantial effort has gone into developing, upgrading, and extending documentation about the SUMEX-AIM resource, the SUMEX-TENEX system, the many subsystems available to users, and MAINSAIL. These efforts include a number of major documents (such as SOS, PUB, and TENEX-SAIL manuals) as well as a much larger number of document upgrades, user information and introductory notes, an ARPANET Resource Handbook entry, and policy guidelines.

3 RESOURCE FINANCES

3.1 BUDGETARY MATERIALS

The budget for the SUMEX project detailing past actual costs, current year status, and estimates for the next grant year are submitted in a separate document to the NIH.

3.2 RESOURCE FUNDING

The SUMEX-AIM resource is essentially wholly funded by the Biotechnology Resources Program (5). The various collaborator projects which use SUMEX are independently funded with respect to their manpower and operating expenses. They obtain from SUMEX, without charge, access to the computing and, in most cases, communications facilities in exchange for the participation in the scientific and community building goals of SUMEX.

---

(5) Except for participation by Stanford University in accordance with general cost-sharing and for assistance to SUMEX from other projects with overlapping aims and interests.

4 COLLABORATIVE PROJECT REPORTS

The following subsections report on the collaborative use of the SUMEX facility. Descriptions are included for the formally authorized projects within the national AIM and Stanford aliquots and the various "pilot" efforts currently under way. These project descriptions and comments are the result of a solicitation for contributions sent to each of the project Principal Investigators requesting the following information:

I. SUMMARY OF RESEARCH PROGRAM
- A. Technical goals
- B. Medical relevance and collaboration
- C. Progress summary
- D. List of relevant publications
- E. Funding support status (see below for details)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE
- A. Collaborations and medical use of programs via SUMEX
- B. Sharing and interactions with other SUMEX-AIM projects (via workshops, resource facilities, personal contacts, etc.)
- C. Critique of resource management (community facilitation and computer services)

III. RESEARCH PLANS (8/78 - 7/81)
- A. Long range project goals and plans
- B. Justification and requirements for continued SUMEX use [This section will be of special importance to the Advisory Committee and constitutes your application for continued access.]
- C. Your needs and plans for other computational resources, beyond SUMEX/AIM
- D. Recommendations for future community and resource development

We believe that the reports of the individual projects speak for themselves as rationales for participation; in any case the reports are recorded as submitted and are the responsibility of the indicated project leaders.

4.1 NATIONAL AIM PROJECTS

The following group of projects is formally approved for access to the AIM aliquot of the SUMEX-AIM resource. Their access is based on review by the AIM Advisory Group and approval by the AIM Executive Committee.

4.1.1 ACQUISITION OF COGNITIVE PROCEDURES

ACQUISITION OF COGNITIVE PROCEDURES (ACT)

Dr. John Anderson
Yale University

I. Summary of Research Program

A. Technical goals:

To develop a production system that will serve as an interpreter of the active portion of an associative network. To model a range of cognitive tasks including memory tasks, inferential reasoning, language processing, and problem solving. To develop an induction system capable of acquiring cognitive procedures with a special emphasis on language acquisition.

B. Medical relevance and collaboration:

1. The ACT model is a general model of cognition. It provides a useful model of the development of and performance of the sorts of decision making that occur in medicine.

2. The ACT model also represents basic work in AI. It is in part an attempt to develop a self-organizing intelligent system. As such it is relevant to the goal of development of intelligent artificial aids in medicine.

We have been evolving a collaborative relationship with James Greeno and Allan Lesgold at the University of Pittsburgh. They are applying ACT to modeling the acquisition of reading and problem solving skills. We plan to make ACT a guest system within SUMEX. ACT is currently at the state where it can be shipped to other INTERLISP facilities. We have received a number of inquiries about the ACT system. ACT is a system in a continual state of development but we periodically freeze versions of ACT which we maintain and make available to the national AI community.

C. Progress and accomplishments:

ACT provides a uniform set of theoretical mechanisms to model such aspects of human cognition as memory, inferential processes, language processing, and problem solving. ACT's knowledge base consists of two components, a propositional component and a procedural component. The propositional component is provided by an associative network encoding a set of facts known about the world. This provides the system's semantic memory. The procedural component consists of a set of productions which operate on the associative network. ACT's production system is considerably different than many of the other currently available systems (e.g., Newell's PSG). These differences have been introduced in order to create a system that will operate on an associative network and in order to accurately model certain aspects of human cognition.

A small portion of the semantic network is active at any point in time. Productions can only inspect that portion of the network which is active at the particular time. This restriction to the active portion of the network provides a means to focus the ACT system in a large data base of facts. Activation can spread down network paths from active nodes to activate new nodes and links. To prevent activation from growing continuously there is a dampening process which periodically deactivates all but a select few nodes. The condition of a production specifies that certain features be true of the active portion of the network. The action of a production specifies that certain changes be made to the network. Each production can be conceived of as an independent "demon." Its purpose is to see if the network configuration specified in its condition is satisfied in the active portion of memory. If it is, the production will execute and cause changes to memory. In so doing it can allow or disallow other productions which are looking for their conditions to be satisfied. Both the spread of activation and the selection of productions are parallel processes whose rates are controlled by "strengths" of network links and individual productions. An important aspect of this parallelism is that it is possible for multiple productions to be applied in a cycle. Much of the early work on the ACT system was focused on developing computational devices to reflect the operation of parallel, strength-controlled processes and working out the logic for creating functioning systems in such a computational medium.

We have successfully implemented a number of small-scale systems that model various psychological tasks in the domain of memory, language processing, and inferential reasoning. There was a larger scale project to model the language processing mechanisms of a young child. This includes implementation of a production system to analyze linguistic input, make inferences, ask and answer questions, etc.

The current research is focused on developing mechanisms for the acquisition of skills. In the framework of the ACT system this maps into acquiring new productions and modifying old productions. We have developed learning devices to enable existing productions to create new productions, to adjust the strengths of existing productions, to produce more general variants of existing productions, to produce more discriminant variants of existing productions, and to combine a number of existing productions into a single compact production. We have developed the F version of the ACT system which has these learning facilities. We have so far tested out the system in a number of small learning examples. Current goals involve applying the system to the acquisition of language skills, development of mathematical problem solving skills, and acquisition of initial programming skills.

The basic insight in this research is to model skill acquisition as an interaction between deliberate learning and automatic induction. To the extent that the teacher or the learner is able to understand the skill to be acquired, it is possible for ACT to directly create the necessary productions. However, as a fallback for less structured situations, ACT has automatic induction mechanisms that try to develop the necessary mechanisms by an intelligent trial and error inductive process. Much of our research has gone to identifying the heuristics used by this inductive process. Traditionally, there has been a contrast in psychology between learning with understanding and learning by trial and error. It is now clear to us that most real learning situations involve a mixture and the key to understanding skill acquisition is to understand that mixture.

D. Current list of project publications:

[1] Anderson, J.R. Language, Memory, and Thought. Hillsdale, N.J.: L. Erlbaum, Assoc., 1976.

[2] Kline, P.J. & Anderson, J.P. The ACTE User's Manual, 1976.

[3] Anderson, J.R., Kline, P. & Lewis, C. Language processing by production systems. In P. Carpenter and M. Just (Eds.). Cognitive Processes in Comprehension. L. Erlbaum Assoc., 1977.

[4] Anderson, J.R. Induction of augmented transition networks. Cognitive Science, 1977, 125-157.

[5] Anderson, J.R. & Kline, P. Design of a production system. Paper presented at the Workshop on Pattern-Directed Inference Systems, Hawaii, May 23-27, 1977.

[6] Anderson, J.R. Computer simulation of a language acquisition system: A second report. In D. LaBerge and S.J. Samuels (Eds.). Perception and Comprehension. Hillsdale, N.J.: L. Erlbaum Assoc., 1978.

[7] Anderson, J.R., Kline, P.J., & Beasley, C.M. A theory of the acquisition of cognitive skills. In G.H. Bower (Ed.). Learning and Motivation, Vol. 13. New York: Academic Press, 1979.

[8] Anderson, J.R., Kline, P.J., & Beasley, C.M. Complex Learning. In R. Snow, P.A. Frederico, & W. Montague (Eds.). Aptitude, Learning, an Instruction: Cognitive Processes Analyses. Hillsdale, N.J.: Lawrence Erlbaum Assoc., 1979.

E. Funding:

ONR Contract N0014-77-6-0242 (total period 3/77 - 9/80)
A Model for Procedural Learning
$90,000 3/78 - 2/79 (including indirect costs)

II. Interaction With the SUMEX-AIM Resource

A. & B. Collaborations, interactions, and sharing of programs via SUMEX.

We have received and answered many inquiries about the ACT system over the ARPANET. This involves sending documentations, papers, and copies of programs. We have also used the ARPANET to access and experiment with the production systems at Carnegie Mellon University. The most extensive collaboration has been with Greeno and Lesgold who are also on SUMEX. There is an ongoing effort to help them in their research. Feedback from their work is helping us with system design.

We find the SUMEX-AIM workshops ideal vehicles for updating ourselves on the field and for getting to talk to colleagues about aspects of their work of importance to us.

C. Critique of resource management.

The SUMEX-AIM resource is superbly suited for the needs of our project. We have made the most extensive use of the INTERLISP facilities and the facilities for communication on the ARPANET. We have found the SUMEX personnel extremely helpful both in terms of responding to our immediate emergencies and in providing advice helpful to the long-range progress of the project. Despite the fact that we are on the other side of the continent, we have felt almost no degradation in our ability to do research. We find we can easily list on the terminal a small portion of programs under modification. The willingness of SUMEX to mail listings has also meant we can keep relatively up-to-date records of all programs.

A unique east coast advantage of working with SUMEX is the low loading of the system during the mornings. We have been able to get a great deal of work done during these hours and try to save our computer-intensive work for these hours.

A particularly striking example of the utility of the SUMEX resource was illustrated in the move from Michigan. In the summer of 1976 Anderson moved to Yale and Greeno to Pittsburgh. There was no loss at all associated with having to transfer programs from one system to another. At Yale we were programming the day after we arrived. The SUMEX link has also permitted continued collaboration with Greeno. We are planning a permanent move to Carnegie-Mellon this summer and happily anticipate it will be as painless.

III. Research Plans (8/78-7/81)

A. Long-range user project goals and plans:

Our long-range goals are: (1) Continued development of the ACT system; (2) Application of the system to modeling of various cognitive processes; (3) Dissemination of the ACT system to the national AI community.

1. System Development We have completed the F version of the ACT system. We are currently applying or intend to apply the ACT system to modeling the acquisition and/or performance of cognitive skills in the areas of language comprehension and generation, inferential reasoning, reading skills, mathematical problem solving, and computer programming. It is hard to anticipate now all the impact of these explorations for design decisions in later versions of ACT. However, it is clear even now that a number of developments are needed. We want to make ACT more appropriate as a language for programming cognitive skills. This involves such things as development of more powerful control conventions, simplification of syntax, and introduction of direct programming features (such as comparison of quantity magnitudes) that can only be obtained indirectly in ACTF. We also want to introduce more efficient implementation techniques to replace some of the simple devices that were used to enable us to rapidly complete the system. These rearchitecture efforts have to be done within the

constraints of psychological plausibility, but we have a theoretical commitment to the conjecture that good implementation design is predictive of good psychological mechanisms. We are currently implementing a new system--G version--which will incorporate these ideas and any additional insights that will come out of our experimentation with ACTF.

2. Application to Modeling Cognitive Processes. We anticipate a gradual decrease in the amount of effort that will go into system development and an increase in the amount of effort that will go into application of the system for modeling. We mentioned above the modeling efforts that we are using to assess the suitability of the ACTF system. We have long-range commitments to apply the ACT learning model to the following three topics: Acquisition of language (both first and second language acquisition); acquisition of programming skills; acquisition of problem solving skills in the domain of geometry. We find each of these topics to be considerable interest in and of themselves, but they also will serve as strong tests of the learning model. We are hopeful that the systems that are acquired by ACT will satisfy computational standards of good artificial intelligence. Therefore, in future years we would also be interested in applying the ACT model to acquisition of cognitive skills in medically related domains such as diagnosis or scientific inference. SUMEX would be an ideal location for collaboration on such a project.

3. Dissemination of the ACT Project We have a commitment to making the ACT system available to anyone in the national community who has access to the necessary computer resources. This is partially to provide a service in that ACT is a medium for psychological modeling. However, it is also self-serving in that the use of other people make of ACT has important feedback in assessing design decisions. In light of limitations on the SUMEX resource, we have decided not to allow extensive use of ACT by other researchers through our SUMEX account. We feel that extensive use of the ACT system in SUMEX by another researcher must have the status of an independent project and must be able to justify independently its use of the SUMEX-AIM resource. The current system being supported for use by other researchers is ACTE but we are in the process of updating our supported system to ACTF.

B. Justification for continued use of SUMEX:

We feel that the justification for our use of SUMEX has only been strengthened since the time of our original application for user status. The project meets a number of criteria for SUMEX relevance: The project is concerned with cognitive modeling which is a SUMEX goal. The project is also developing an AI tool which can be used to help automate various medically-relevant tasks. We also think we are the type of need that the SUMEX facility was designed to meet. That is, we do not have nearly as powerful computing facilities local at Yale; we are non-local user; we are using SUMEX as a base for collaborating with scientists in other parts of the country; and we are trying to develop a system that will be of general use.

Our future move to Carnegie Mellon raises some interesting issues about the SUMEX resources. The availability of the SUMEX resource makes the move easy and allows the research to go full steam ahead. The fact that Carnegie-Mellon is on the ARPANET will reduce our cost (no TYMNET charges) and allows us to get immediate listings which had been the one deficit of being a distant user. On the other hand, the greater Carnegie resources may diminish the need for SUMEX.

We have not carefully explored the quality of the Carnegie system versus SUMEX but an important obstacle for us is the lack of INTERLISP at Carnegie. Carnegie will have INTERLISP available within the next two years. At such time it may be appropriate to enter discussion with SUMEX about how to balance our lesser needs for SUMEX with our cost to SUMEX and both of these factors with the role we play in the SUMEX community.

C. Comments and suggestions for future resource goals:

We would, of course, be delighted if the computational capacity of the SUMEX facility could be increased. The slowness of the system at peak hours is a limiting factor although it is not grievous. This problem is perhaps less grievous for us than Stanford-based users because of our ability to use morning hours. We do not feel any urgent need for development of new software.

4.1.2 CHEMICAL SYNTHESIS PROJECT

SECS - Simulation and Evaluation of Chemical Synthesis

Principal Investigator: W. Todd Wipke
Board of Studies in Chemistry
University of California at Santa Cruz

Coworkers: S. Krishnan, C. Buse, and M. Huber (Postdoctoral Fellows)
G. Ouchi and D. Dolata (Grad students)

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The long range goal of this project is to develop the logical principles of molecular construction and to use these in developing practical computer programs to assist investigators in designing stereospecific syntheses of complex bio-organic molecules. Our specific goals this past year focused on strategic control of the SECS program, on implementing strategies based on symmetry and potential symmetry, on developing ways to treat the steric factors involved in acyclic reaction centers, on converting SECS from F40 FORTRAN to F10 FORTRAN (the current supported FORTRAN), and on demonstrating that computer synthesis techniques are also applicable to metabolism. In addition, we wanted to add a library lookup capability, and to improve the user interaction with the SECS program.

B. Medical Relevance and Collaboration.

The development of new drugs and the study of how drug structure is related to biological activity depends upon the chemist's ability to synthesize new molecules as well as his ability to modify existing structures, e.g., incorporating isotopic labels or other substituents into biomolecular substrates. The Simulation and Evaluation of Chemical Synthesis (SECS) project aims at assisting the synthetic chemist in designing stereospecific syntheses of biologically important molecules. The advantages of this computer approach over normal manual approaches are many: 1) greater speed in designing a synthesis; 2) freedom from bias of past experience and past solutions; 3) thorough consideration of all possible syntheses using a more extensive library of chemical reactions than any individual person can remember; 4) greater capability of the computer to deal with the many structures which result; and 6) capability of computer to see molecules in graph theoretical sense, free from bias of 2-D projection.

The objective of using SECS in metabolism is to predict the plausible metabolites of a given xenobiotic in order that they may be analyzed for possible carcinogenicity. Metabolism research may also find this useful in the identification of metabolites in that it suggests what to look for. Finally, it seems there may even be application of this technique in problem domains where one wishes to alter molecules so certain types of metabolism will be blocked.

C. Progress and Accomplishments

RESEARCH ENVIRONMENT: At the University of California, Santa Cruz, we have a GT40 graphics terminal connected to the SUMEX-AIM resource by a 1200 baud leased line (the leased line supported by SUMEX). On 1 May 1978 a new GT46 graphics terminal was installed in addition. We also have a TI725, a TI745, a CDI-1030, a DIABLO 1620, and an ADM-3A terminal which were used over the University tie-line system to SUMEX, but now are awaiting leased lines since the University of California did away with the tie-line system. UCSC has only a small IBM 370/145, a PDP-11/45 and 11/70 (the latter are limited to small student time-sharing jobs of 12 K words per user), all of which are unsuitable for this research. We hope through the GT46 system, to be able to transfer files from SUMEX to the UCSC PDP-11 for local printing, and possibly local magnetic tape handling (currently all input and output except graphical must be done at Stanford.) The SECS laboratory is located in the same building as the synthetic chemists at Santa Cruz so there is very facile interaction.

STRATEGIC CONTROL: We feel, and feedback from users supports this, that there is a great need for the user of SECS to control and direct the synthesis planning, if the user so desires. For the purpose of this discussion, a strategy is a general principle which helps guide one in generating a simple synthesis. Strategies are based on symmetry, mathematical considerations of yield, economy of operations, etc. When a strategy is applied to a particular synthetic target molecule, it generates goals. Goals are described only in terms of molecular structural changes or features, and may not, for example, refer to reactions. Thus, strategies create goals, and both are completely independent of the reaction library.

We had previously created a list-structured language for describing goals to allow manual introduction of goals which can then direct selection of relevant transforms to only those that satisfy the specified goals. We have continued to improve the human interface to that module, but the majority of our current work is on developing modules which implement various high level strategies and thus automatically create goals. The chemist of course will still have the opportunity to modify these automatically created goals. The reason for wanting to create them automatically is that the chemist would never be able to consider all possible strategies and much of the creativity in a synthesis is in selection of the basic strategy and resulting goals. And the reason for having goals is that it gives SECS a sense of purpose and justifies therefore devoting more resources to following those goals and trying to achieve them, sometimes to the exclusion of other chemical transformations. The power of the goal list has led to some unexpected capabilities. For example, by specifying a certain set of changes in the target molecule and that only those transforms which satisfy that goal list are allowed to be applied, one can find out whether SECS has a reaction in its library capable of making those changes. The following paragraphs describe some of the current strategy work.

STRATEGIES BASED ON SYMMETRY: Based on analyses of many literature syntheses, we have found several key strategies related to symmetry. One of these involves trying to break the structure into two or more identical fragments. The advantages of a synthesis utilizing identical or similar fragments result both from a minimization of the number of synthetic steps, and from the principle of convergent synthesis (if the identical fragments constitute a major part of the target compound). This analysis takes place on the strategic level, meaning that SECS considers the non-redundant disconnections which generate fragments, regardless of whether any applicable reactions exist which could actually form the bond broken in the analysis. Identical fragments are recognized by comparing their SEMA canonical names. In order to find similar fragments we first had to define what "similar" meant and then to define a function to calculate a metric on that space. The current function considers atom types, bond types, and stereochemistry in addition to connectivity, and it includes weighting factors for each term. At the end of the analysis, fragment sets containing identical or similar fragments are shown to the user and can then be converted into GOALS. Later, SECS uses those GOALS to find transforms that can achieve the desired constructions.

When molecules that have been synthesized from identical or similar fragments (literature) are analyzed by the program, it so far has always been able to detect the proper fragmentations. Interestingly, in analyzing natural products, SECS "discovered" the isoprene rule since nature uses this strategy too! This module generates some very challenging suggestions, even though there may not be known reactions to implement the suggestions. In fact, this now provides SECS with a rationale which in the future might suggest a need for a new reaction.

SUBGOALS: When a chemical transform has a high priority and seems to be able to satisfy a goal on the goal list the transform is "relevant", but still may not be "applicable" owing to some mismatch between what the transform requires and what the operand structure has. This mismatch can spawn a SUBGOAL to change the structure until this transform is applicable. The first utilization of subgoals in SECS is for automatic functional group interchange (FGI). Mismatches in the identity of functional groups are easier to correct than mismatches in the carbon skeleton. The program now recognizes these functional group mismatches, generates subgoals, tries to satisfy the subgoals, and then goes back to the original goal. This is equivalent to a "look-ahead" and it may lead to sequences of several reactions, but SECS invents the sequences, they are not preprogrammed. At the time of this writing, this module works well on small molecules, but on complex ones, the number of subgoals created is too large. Methods of combating this problem are under consideration.

STRUCTURE DICTIONARY: "Is this compound commercially available?", "Is this a new compound?", and "Is this a known carcinogen?" are some of the questions a synthetic chemist asks in designing a synthesis. In order to be able to answer questions like these rapidly regardless of how large the chemical dictionary may be, we developed an efficient search method using a hash of the SEMA canonical name. We studied the efficiency of several hash functions, with various sized tables and varying amounts of detail in the SEMA name. On a typical file of compounds from a synthesis, our hash function randomized the keys as well as a random number generator. SECS uses this technique now in finding if a precursor already exists in the synthetic tree. We have also used this for finding common biosynthetic intermediates by checking the precursor against other "trees" which were generated in previous analysis of other structures. We plan to add other libraries of compounds in the future.

STEREOCHEMICAL INDUCTION: In the synthesis of macrocyclic natural products such as erythronolide A, or maytansine, stereochemical control is of prime importance. This can be achieved by using small rings, or by using stereospecific acyclic carbonyl addition reactions which follow Cram's rule. We have implemented a module to evaluate Cram's rule and alter the reaction's priority value accordingly. The module has been tested using over 50 reactions from the literature.

EXPLANATION CAPABILITY: We feel that in order to gain the confidence of the user, SECS should be able to explain the source of its knowledge and some of its reasoning. Our first step in this direction is a rapid retrieval of literature references for transforms. The second step is modeled after the question-answering of MYCIN. For some time now, SECS has recorded the types of questions chemists ask. A parser for those questions is now underway with the answer generator to follow. We hope to be able to answer questions such as "Why did you not use the Wittig reaction here?" or "Why is the priority of this reaction so high?"

METABOLISM PREDICTION: Numerous structurally different chemical compounds have been found to induce neoplasia in man and animals. In many cases these chemical carcinogens are metabolically activated by mammalian enzyme systems to their ultimate reactive and toxic structure. Many of the mechanisms involved in this "bioactivation" process are known or are in the process of being discovered. Thus, it is now possible based on the structure of a compound and a through knowledge of biotransformations to make rational predictions of the plausible metabolites of a compounds produced in a mammalian system. To study the metabolic activation of compounds we are creating a computer assistant which will generate the plausible metabolites of a compound utilizing the biotransformations known to occur in mammalian systems.

A new computer program called XENO for the metabolism of xenobiotic compounds has been developed based on technology from computer synthesis project. However, since metabolism is being simulated in the forward direction, whereas organic synthesis is simulated in the reverse direction, the XENO program is quite different in logic from SECS, although both use ALCHEM as a representation for reactions. The XENO data base of biotransforms was developed by careful survey of metabolism literature and consultation with a committee of metabolism experts at NIH. We selected a mechanistic representation of metabolic processes which means a small data base suffices to represent most of the known processes. A critical evaluation of XENO by a panel of experts in Bethesda, Md. in February 1978 concluded that the data base of biotransforms must be considerably expanded, but even now it is able to raise some interesting questions of alternative metabolic pathways, etc. XENO is currently running on SUMEX-AIM. Shown below is part of the analysis XENO performed on benzo(a)pyrene.

```
:TREE
   1:  2   3   6   9  12  13  14  15  16  17  19  24  26
      27  28  29  30  31  32  33  34  35  36  37  38  39
   2: 40  41  44  46  47  48  52  53
  40: 54  55  56  57  58  59  61  66  68  69  70  71  72
      73  74  75  76  77  78  79  80  81  82  83  85  86
      87  88  89  90  91  92  93  94  95  96
```

1

```
      C=C
     /   \
  C-C     C=C
 =   =   /   \
C     C-C     C-C
 \   /   =   =   =
  C=C     C-C     C
     \   /   \   /
      C=C     C=C
```

Benzo(a)pyrene

2

```
      C=C
     /   \
  C-C     C=C     O
 =   =   /   \   /|
C     C-C     C-C |
 \   /   =   =   \|
  C=C     C-C     C
     \   /   \   /
      C=C     C=C
```

Priority = 120

40

```
          C-C
         =   =
      C-C     C
     =   \   /
    C     C=C
     \   /   \
      C=C     C
     /   \   =
  C-C     C-C
 =   =   =
C     C-C
 \   /
  C-C
 /   \
O     O
```

Priority = 105

54

```
          C-C
         =   =
      C-C     C
     =   \   /
    C     C=C
     \   /   \
O     C=C     C
|\   /   \   =
| C-C     C-C
|/   =   =
C     C-C
 \   /
  C-C
 /   \
O     O
```

Priority = 240 *

55

```
          C-C
         =   =
      C-C     C
     =   \   /
    C     C=C
     \   /   \
O     C=C     C
|\   /   \   =
| C-C     C-C
|/   =   =
C     C-C
 \   /
  C-C
 /   \
O     O
```

Priority = 140 *

D. List of Current Project Publications

S.A. Godleski, P.v.R. Schleyer, E. Osawa, and W.T. Wipke, "The Systematic Prediction of the Most Stable Neutral Hydrocarbon Isomer," J. Am. Chem. Soc., 99, 0000(1978).

W. T. Wipke, G. Smith, F. Choplin, and W. Sieber, "SECS--Simulation and Evaluation of Chemical Synthesis: Strategy and Planning," ACS Symposium Series No. 61, 97-127 (1977).

W.T. Wipke, "Computer Planning of Research in Organic Chemistry," in Computers in Chemical Education and Research, ed. E.V. Ludena, N.H. Sabelli, and A.C. Wahl, Plenum Press, N.Y., 1977, pp. 381-391.

W.T. Wipke, S. Krishnan, and G.I. Ouchi, "Hash Functions for Rapid Storage and Retrieval of Chemical Structures," J. Chem. Info. and Computer Sci., 18, 32 (1978).

F. Choplin, R. Marc, G. Kaufmann, and W.T. Wipke, "Computer Design of Synthesis in Phosphorus Chemistry. Automatic Treatment of Stereochemistry," J. Chem. Info. and Computer Sci., 18, 000 (1978).

F. Choplin, R. Dorschner, G. Kaufmann, and W. T. Wipke, "Computer Graphics Determination and Display of Stereoisomers in Coordination Compounds," in press.

F. Choplin, C. Laurenco, R. Marc, G. Kaufmann, and W.T. Wipke, "Synthese Assistee par Ordinateur en Chimie des Composes Organophosphores," Nouveau J. de Chimie, in press.

W.T. Wipke, G. Ouchi, and S.Krishnan, "Simulation and Evaluation of Chemical Synthesis - SECS. An Application of Artificial Intelligence Techniques," Artificial Intelligence, in press.

M. Spann, K. Chu, W.T. Wipke, and G. Ouchi, "Computer-Aided Prediction of Metabolites," in press.

E. Funding Status

1. Resource-Related Research: Biomolecular Synthesis
   PI: W. Todd Wipke, Associate Professor, UCSC
   Agency: NIH, Research Resources
   No: RR01059-01
   7/1/77-6/30/80 $227,816 TDC
   7/1/77-6/30/78 $94,602 TDC

2. Computer-Aided Prediction of Metabolites for Carcinogenicity Studies
   PI: W. Todd Wipke
   Agency: NIH, National Cancer Institute
   No: N01-cp-75816
   3/8/77-9/7/78 $44,146 TDC
   3/8/78-9/7/78 $14,607 TDC

3. Computer Synthesis, Unrestricted
   PI: W. Todd Wipke, Associate Professor, UCSC
   Agency: Sandoz, Ltd. $2500
   Agency: Bayer $5000
   Agency: E. Merck $1500

II. INTERACTIONS WITH SUMEX-AIM RESOURCE

A. Collaborations and Medical Use of Programs via SUMEX

SECS is available in the GUEST area of SUMEX for casual users, and in the SECS DEMO area for serious collaborators who plan to use a significant amount of time and need to save the synthesis tree generated. Much of the access by others has been through the terminal equipment at Santa Cruz because graphic terminals make it so much more convenient for structure input and output. We have assisted Professor J.E. McMurry of UCSC in his synthetic work towards aphidicoline and digitoxigenin (Total Synthesis of Cardiac Aglycones, HL-18118) using the model builder of SECS for evaluating plausible modes of ring closure. Numerous visitors to UC Santa Cruz have tried their own problems on the SECS program, generally taking away at least a couple of new ideas for research. Professor Ken Williamson of Mt. Holyoke College has made arrangements to access SECS to obtain structures for C-13 nmr analysis, and a student at the University of Mass. Amherst has made arrangements to do several analyses on SECS as an independent research project, the results to be tested in the laboratory. The entire collaboration between Drs. Ted Gram of Guarino's lab, Lance Pohl from Gillette's lab, Dhiren Thakken and Harukiko Hagi from Jerina's lab, Ken Chu and Sid Siegel (chemical carcinogenesis), and Mel Spann (National Library of Medicine) in Bethesda would not be possible without access to XENO through SUMEX.

Synthetic chemists are beginning to come to us for a SECS analysis before beginning a laboratory synthesis. Dr. McMurry for example did a rather complete analysis of morphine before launching his recently successful synthesis. We have also collaborated in the biogenesis work with Professor Phil Crews (UCSC) in marine natural product biogenesis. Dr. Wipke has also used several SUMEX programs such as CONGEN in his course on Computers and Information Processing in Chemistry.

B. Examples of Sharing, Contacts and Cross-fertilization with other SUMEX-AIM projects

Dr. Wipke spent the Winter Quarter on sabbatical at Stanford and regularly attended the SIGLUNCH seminars of the Heuristic Programming Project. We have had several discussions with the MYCIN group about our interest in an explanation capability for SECS. The AIM conference at Rutgers each year has been extremely valuable in generating ideas of new ways to apply current developments in AI to the problem of organic synthesis. Finally, it is impossible to count the daily exchanges that occur between researchers in the SECS group and other members of the AIM community on things related to languages, conferences, papers, seminars, and program sharing. Now that our GT46 is installed, for example, we have been communicating with Achenbach at Stanford regarding the AMOK file transfer system which will help us get local printing of files.

C. Critique of Resource Services

We find the SUMEX-AIM network very well human engineered and the staff very friendly and helpful. The SECS project is probably one of the few on the AIM network which must depend exclusively on remote computers, and we have been able to work rather effectively via SUMEX. Basically we have found that SUMEX-AIM provides a productive and scientifically stimulating environment and we are thankful that we are able to access the resource and participate in its activities. SUMEX-AIM gives us at UCSC, a small university, the advantages of a larger group of colleagues, and interaction with people all over the country. We especially thank SUMEX for support of the leased line for our GT40.

D. Collaborations and Medical Use of Programs via Computers other than SUMEX

Arrangements between the University of California, Santa Cruz and NIH have been begun to try to install a version of SECS on the NIH PDP-10 computer system, and possibly later on the NIH-CIS system. Under an arrangement approved in 1974 between First Data, Princeton University, and NIH, SECS has been available over TELENET so that the public could evaluate the state of the technology first hand, by simply contacting First Data. First Data was selected because that is the system the NIH PROPHET program is also on. As a result of that arrangement, anyone who wishes can use the SECS program without worrying about converting code for their machine, and a number of people in the private sector both in the US and abroad have done so. Beryl Dominy of Pfizer wrote a paper about his experiences with the system.

The AIM-Executive committee has details of all developmental efforts on SECS in the United States and abroad as well as the plans for export of the SECS program. The University of California is handling arrangements for pre-release of the SECS program and will also handle final release of SECS, and is assuring that no conflicts of interest arise according to the NIH and University of California policies.

III. <u>RESEARCH PLANS (8/78-7/81)</u>

A. Long Range Project Goals and Plans

The SECS project now consists of two major efforts, computer synthesis and metabolism, the latter being a very young project. The SECS program with automatic FGI's and strategic controls that are now nearly useable represents a version which we feel will be satisfactory for distribution. Consequently, considerable effort over the next year will be toward that goal. We plan to continue exploring new aspects of the computer synthesis problem, such as depth first analysis, forward simulation of synthesis, and bidirectional search from target to key intermediate. There are many high level strategies of synthesis to be represented. The explanation capability, starting material libraries and starting material strategies, etc, are important features to be included. We have many ideas on how the current model builder could be made more general, faster, and more useful. A rewritten, smaller version of SECS which had an excellent TTY input and output module would be useful for possible installation on the NIH-CIS system where memory and cpu time are very critical to users.

The XENO metabolism project will be expanding the data base to cover more metabolic transforms, including species differences, sequences of transforms, and stereochemical specificities of enzymatic systems. A second phase will apply our "similarity" function to determine when metabolites are similar to known carcinogens. We are also hoping to develop programs which will help maintain the growing data bases. It is not clear at this time how quantitative we can hope to be with XENO's predictions and that will be studied.

Considerable development must occur to fully utilize the capabilities of the new GT46, including file transfer, and improvements to the graphics monitor to allow more local computations on the GT46 and GT40. We intend to keep the SECS program compatible with the entire GT40 family of terminals as it now is. We will be working on various types of hardcopy output, for the synthesis tree, sequences of reactions, the contents of libraries, and possibly even chemical transforms. In the next year we expect to see the SECS program used by more academic researchers over SUMEX than previously and they will need ways to obtain hardcopy of their results.

B. Justification and Requirements for Continued use of SUMEX

The SECS and XENO projects require a large interactive time-sharing capability with high level languages and support programs. I am on the campus computing advisory committee and know that the UCSC campus is not likely in the future to be able to provide this kind of resource. Further there does not appear to be in the offing anywhere in the UC system a computer which would be able to offer the capabilities we need. Thus from a practical standpoint, the SECS and XENO projects still need access to SUMEX for survival. Scientifically, interaction with the SUMEX community is still extremely important to my research, and will continue to be so because of the direction and orientation of our projects. Collaborations on the metabolism project and the synthesis project need the networking capability of SUMEX-AIM, for we are and will continue to be interacting with synthetic chemists at distant sites and metabolism experts at the National Cancer Institute. Our requirements are for good support of FORTRAN compiler and linker, particularly the capability to overlay our programs even on TENEX because in the outside community there is a need to overlay and we want to be able to assure that our program structure is overlayable, and in fact, we must soon overlay SECS on TENEX. Currently there is a bug in the loader which requires one to patch the SAV file every time one loads the program. We will also need ports into SUMEX since we will be using leased lines rather than TYMNET unless TYMNET installs a node in Santa Cruz.

We still are hopeful that someday there will be a good file transfer between TENEX and DECSYSTEM-10 machines via magnetic tape. That is still a difficult problem for anyone who is communicating with the outside world via tape.

C. Needs beyond SUMEX-AIM

Our needs are basically to be able to get local printing and solve the magtape file problems since we envision having to send out a great number of tapes to interested users.

D. Recommendations for community and resource development

The AIM workshop is excellent, particularly if it is held on the WEST COAST once in a while. From a chemistry standpoint, the joint group meetings with the DENDRAL group plus ability to attend seminars at Stanford really satisfy our needs for communication with people of similar interests. We have proposed a workshop for the benefit of the implementors rather than the principal investigators and administrators, for that would do wonders to develop the human resource. We feel the computer resource is rather efficiently used right now. At some point it may make sense to have a chemically oriented system as those applications develop.

4.1.3 HIGHER MENTAL FUNCTIONS PROJECT

MODELING OF HIGHER MENTAL FUNCTIONS

Kenneth Mark Colby, M.D.
Professor of Psychiatry, Biobehavioral
Sciences and Computer Science
University of California at Los Angeles

I. Summary of Research Program

A. Technical Goals:

The goals of this project lie in three areas of research:

(1) Improvement of and experimentation with a computer simulation of paranoid processes (PARRY). Little is known about the optimal treatment of paranoid conditions in psychiatry. Experiments with the model lead to specific recommendations for the communicative therapy of paranoid states and reactions.

(2) The development of a new psychiatric taxonomy for incapacitating neuroses of adulthood. The current classification scheme in this area is full of uncertainties and unreliabilities. A more reliable taxonomy would permit specific treatment interventions to be applied to more homogeneous groups of patients.

(3) The construction and development of an intelligent speech prosthesis for patients with language disorders, especially those who have suffered a stroke. Stroke is a double disaster because the patient not only suffers a serious illness but also may be unable to communicate with anyone about it or about himself. An intelligent speech prosthesis would provide a means of restoring spoken communication between the patient and those around him.

B. Medical Relevance and Collaboration:

This project is located in the UCLA Neuropsychiatric Institute. The problems of paranoia, psychiatric classification and language disorders secondary to stroke are of obvious medical relevance. The project members collaborate with a large number of psychiatrists, psychologists, neurolinguists and computer scientists working in biobehavior and medical fields.

C. Progress Summary:

In the past year we have developed an improved version of the paranoid model which now "thinks" by itself and treats input as an interrupt condition. The major improvement lies in its ability to keep better track of the conversation.

In establishing a Clinical Mental Health Research Center at UCLA we have been collaborating with seven other investigators in psychiatry, ethology, electrophysiology, psychophysiology, neuropsychology, psychology and statistics in an effort to construct a new taxonomy for the severe neuroses. This population, while large (10-15% of the total U.S. population), has been little studied and hence is little understood. A reliable taxonomy must be developed as a first step. We are studying these patients from a number of perspectives. The Higher Mental Functions project contributes a cognitive assessment of the patient which attempts to categorize him in terms of his concepts and key ideas. We are just beginning to collect this data and correlating it with data from other assessments. The Principal Investigator of this project is also the PI and Director of the Clinical Mental Health Research Center.

The intelligent speech prosthesis now has about 800 orthographic-to-phonetic rules which give it an infinite vocabulary. We have constructed an ear-button allowing the user to correct errors through auditory as well as visual feedback. We have drawn the schematics for a portable device which can be worn on the arm and are in the process of building this portable device. Also we have begun to construct a dynamic associative network which aids the user in word-finding. This type of semantic and phonetic memory will become unique to the user over time because the device keeps track of which associations are used more often in word-searching.

D. List of Relevant Publications:

Parkison, R. C., Colby, K. M. and Faught, W. S. Conversational Language Comprehension Using Integrated Pattern-Matching and Parsing. ARTIFICIAL INTELLIGENCE, 9, 111-134, 1977.

Colby, K. M. On the Way People and Models Do It. PERSPECTIVES IN BIOLOGY AND MEDICINE, 21, 99-104, 1977.

Colby, K. M., Christinaz, D., Parkison, R. C. and Faught, W. S. A Proposed Psychiatric Taxonomy Based On the Conceptual Patterns and Key Ideas in Patient Accounts. ALGORITHMIC LABORATORY OF HIGHER MENTAL FUNCTIONS PROJECT, MEMO ALHMF-10, 1977. UCLA Department of Psychiatry. Colby, K. M. Mind Models: An Overview of Current Work. MATHEMATICAL BIOSCIENCES, In Press, 1978.

Colby, K. M., Christinaz, D. and Graham, S. A. A Computer-Driven, Personal, Portable, and Intelligent Speech Prosthesis. COMPUTERS AND BIOMEDICAL RESEARCH, June, 1978.

E. Funding Support Status:

This project is funded entirely by intradepartmental funds from the Biobehavioral Sciences Program. No other contracts or grants are at present involved. Applications for grants have been made to the National Institute of Mental Health, the National Science Foundation and the National Institute of Neurological and Communicative Disorders and Stroke.

II. Interactions with the SUMEX-AIM Resource

A. Collaborations

We have collaborated with Professor Jon Heiser, Department of Psychiatry, University of California, Irvine in testing the paranoid model. Also we have consulted with Professor John Eulenberg, Department of Computer Science, Michigan State University, regarding problems of intelligent artificial speech. Many informal consultations also take place with other SUMEX users regarding specific problems.

B. Sharing:

Members of the project participate in the Rutgers AIM workshops and in other conferences which SUMEX users also attend.

C. Criticism:

We have no criticism of SUMEX itself. It is easily the best system we have ever used. Our only criticism is directed at TYMNET which is so unreliable (garbling characters) as to be practically unusable.

III. Research Plans

A. Long Range Goals:

We plan to continue to work on the problems of paranoid conditions, psychiatric taxonomy, and intelligent speech prosthesis over the next three years at least.

B. Justification for Continued SUMEX Use:

Without SUMEX this project could not continue its work on paranoia and on psychiatric taxonomy. Both of these efforts involve large LISP programs which must be run interactively on a large time-shared system. Work on intelligent artificial speech requires use of large dictionaries such as available on SUMEX. We do not anticipate requiring any further space allocations than we now have.

C. Other Computational Resources:

We use microprocessors and a 360/91 at UCLA in addition to SUMEX in the research on intelligent artificial speech.

D. Recommendations:

We strongly recommend that the SUMEX-AIM facility be continued. Otherwise we would not be able to continue the bulk of our research efforts.

4.1.4 INTERNIST PROJECT

INTERNIST - Diagnostic Logic Project

J. Myers, M.D. and H. Pople, Ph.D.
University of Pittsburgh

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The major goal of the INTERNIST project is to produce a reliable and adequately complete diagnostic consultative program in the field of internal medicine. Although this consultative program is designed primarily to aid skilled internists in complicated medical problems, the program may have spin-off as a diagnostic and triage aid to physicians assistants, rural health clinics, military medicine and space travel.

To be effective, the program must be capable of multiple diagnoses (related or independent) in a given patient and it should deal effectively with the time axis in the development and course of disease states.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The knowledge base should reach a critical stage of completeness within a year to 18 months, at which point we shall invite collaboration in the field testing of the program in a number of medical institutions. Desires for such collaboration have been very positively indicated by more than an adequate number of sister academic health centers and community hospitals, etc.

The Department of Pediatrics at Pittsburgh has just begun a collaboration with INTERNIST with the objective of a similar diagnostic program in the field of pediatrics.

C. Progress Summary

The original INTERNIST program described in previous progress reports and documented in Pople, Myers & Miller [3] continues to be the standard diagnostic program used to analyze clinical problems and to exercise newly developed portions of the knowledge base.

The structure of the medical knowledge base has remained comparatively constant during the past year. The knowledge base has been expanded by the addition of some fifty diseases. The existing knowledge base is under a process of continual editing which attempts to keep the data up to date by the addition of new information about diseases as such becomes available, and which expands and corrects the old data base as omissions or errors are discovered. To our gratification, the progressive enlargement of the knowledge base has in no significant adverse way affected the operation of the computer program.

The program and the knowledge base are continually being tested with challenging medical problems with good and reasonable success. The knowledge base remains too incomplete for any comprehensive or critical test on our hospital floors but the system is used on an ad hoc basis for clinical guidance.

Experience with this system has led to the identification of certain performance deficiencies that are being addressed in the design of a second generation diagnostic program (INTERNIST-II) the essential features of which are outlined in Pople [1]. A major objective in the design of the new program is to enable concurrent evaluation of the multiple components of a complex clinical problem, thereby enhancing the system's rate of convergence on the essential nature of the problem. A number of new concepts, not presently captured in the existing INTERNIST knowledge base, are required for this purpose; for example: the "constrictor" relation described in [1]; generalization of the INTERNIST disease hierarchy to a network permitting multiple categorization, and incorporation of "temporal schemata" used to permit expression of the time course of disease processes. The detailed design of these data structures and the diagnostic procedures that use them has progressed to the point of preliminary testing, but it should be emphasized that this is a research area in which considerable work remains to be done.

Two full-time graduate students in computer science and two full-time clinical fellows in medicine will be working on the INTERNIST project in 1978-79. This constitutes our largest training staff to date and these young persons will undoubtedly contribute significantly to the speed and quality of the project's continuing development.

D. Publications

1. Pople, H.E. "The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning", Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Boston, August 1977.

2. Pople, H.E. "On the Knowledge Acquisition Process in Applied A.I. Systems", Report of Panel on Applications of A.I., Proceedings of Fifth International Joint Conference on Artificial Intelligence, 1977.

3. Pople, H.E., Myers, J.D. & Miller, R.A. "The DIALOG Model of Diagnostic Logic and its Use in Internal Medicine, Proceedings of the Fourth International Joint Conference on Artificial Intelligence, Tbilisi, USSR, September 1975.

4. Pople, H.E. "Artificial-Intelligence Approaches to Computer-based Medical Consultation, Proceeding IEEE Intercon, New York, 1975.

E. Funding Support Status

1. Title of Grants:
   (a) Dialog: A Computer Model of Diagnostic Logic
   (b) Clinical Decision Systems Research Resource

2. Principal Investigators:
   (a) Jack D. Myers, M.D. University of Pittsburgh
       Harry E. Pople, Ph.D. University of Pittsburgh
   (b) Jack D. Myers, M.D. University of Pittsburgh
       Harry E. Pople, Ph.D. University of Pittsburgh

3. Funding Agencies:
   (a) BHRD
   (b) NIH

4. Grant or Contract Identification number:
   (a) 5R01 MB00144 04 (total period 7/74 - 6/78)
   (b) 1R24 RR01101 01 (total period 7/77 - 6/80)

5. Current Terms:
   (a) 6/30/77-6/29/78
       $101,000 (direct costs)
   (b) 7/1/77-6/30/78
       $160,000 (direct costs)

II. INTERACTIONS WITH SUMEX-AIM RESOURCE

A,B. Collaborations and Medical Use of Program Via SUMEX

INTERNIST remains in a stage of research and development. As noted in the "Progress Summary" above, we are continuing to attempt to develop better computer programs to operate the diagnostic system, and the knowledge base cannot be used very effectively for collaborative purposes until it has reached a critical stage of completion. These factors have stifled collaboration via SUMEX up to this point and will continue to do so for the next year or two. In the meanwhile, through the SUMEX community there continues to be an exchange of information and states of progress. Such interactions particularly take place at the annual AIM Workshop.

Dr. Victor Yu, formerly associated with MYCIN, is now a faculty member at the University of Pittsburgh and has begun active participation in INTERNIST. Dr. Yu will be particularly valuable in the programming of infectious diseases.

C. Critique of Resource Management

SUMEX has been an excellent resource for the development of INTERNIST. Our large program is handled efficiently, effectively and accurately. The staff at SUMEX have been uniformly supportive, cooperative, and innovative in connection with our project's needs.

III. RESEARCH PLANS (8/78 TO 7/81)

A. Long Range Project Goals and Plans

The primary goal of INTERNIST is to develop and complete an effective and reliable instrument for diagnostic consultation in internal medicine. To accomplish this a very extensive knowledge base must be developed, tested and continually updated. The initial stage of development is about 75% accomplished; a reasonable complete knowledge base, incorporating the new data structures identified in section I above, is a year to 18 months in the future. With this development together with the improvement in the computer analytical program, INTERNIST will be suitable for a critical field trial, first in our own health center and, assuming success, in a half-dozen or so of additional health care institutions. Successful completion of the field test should make the program ready for practical clinical use.

B. Justification and Requirements for SUMEX Use

Neither the continued evaluation and development of INTERNIST's computer program nor the manipulation and further development of INTERNIST's knowledge base can be accomplished without a large computer resource such as SUMEX. SUMEX has thus far met our requirements admirably and those requirements for the research and development component of INTERNIST should remain relatively constant over the next three years. The SUMEX resource (or its equivalent) is absolutely essential to INTERNIST's progress.

C. Needs and Plans for Other Computational Resources

As predicted above, INTERNIST should be ready for field testing within two years. It is realized that it is not the purpose to SUMEX in its present form to support such extensive trials. Accordingly, a dedicated computer (or a dedicated portion of SUMEX) will be needed to carry out the trials. No specific plans have yet been made for this operation.

4.1.5 MEDICAL INFORMATION SYSTEMS LABORATORY

MISL - Medical Information Systems Laboratory

M. Goldberg, M.D. and B. McCormick, Ph.D.
University of Illinois at Chicago Circle

I) SUMMARY OF RESEARCH PROGRAM

A) TECHNICAL GOALS

The Medical Information Systems Laboratory (MISL) was established under grant HM-0114 in Chicago to pursue three activities: i) Construction of a database in ophthalmology, ii) Clinical knowledge system support, and iii) Network-compatible database design. Priorities in year 04 of MISL's operation are the same as in previous years: investigations into how to construct a database in ophthalmology, and into distributed database design, are ancillary to the exploration of a clinical knowledge system to support clinical decision making. We are developing ways to get reliable clinical information into the ophthalmic database primarily because we are interested in getting out significant clinical decision support.

B) APPROACH AND MEDICAL RELEVANCE

B.1) Construction of the database in Ophthalmology

A specific aim of this project is to construct a workable database in ophthalmology, using the outpatient population of the Illinois Eye and Ear Infirmary. We view this database as a testbed for developing clinical decision support systems. The Ophthalmology Department of the Illinois Eye and Ear Infirmary provides an excellent environment for evaluating new techniques for capturing and using clinical information.

B.2) Clinical knowledge support system

The goals for clinical knowledge system development are to provide a flexible user interface for a prototype relational database system, to devise means of accessing alphanumeric and pictorial information stored in the database system, and to provide efficient means for logically restructuring a database so that it can be adapted to different operating environments in a network-compatible distributed medical information network.

C) PROGRESS SUMMARY (INCLUDING ITEMS OF INTEREST TO SUMEX-AIM COMMUNITY ONLY)

C.1) The database in ophthalmology

Since last year, the clinical support computer system has been relocated to the Goldberg Research Center, across a street from the Illinois Eye and Ear Infirmary at the University of Illinois Medical Center Campus. Physician terminals and interfaces to ophthalmic instruments have been positioned in the general eye clinic of several key ophthalmic subspecialty clinics. Systematic, modular hardware and software for clinical source data acquisition have been established. The hardware configuration and telecommunication linkages have stabilized. We are presently using RAIN, a relational database system developed by Prof. S. K. Chang of MISL, to collect general eye index data, along with data about glaucoma.

C.2) Clinical knowledge system support

C.2.a) Development of the relational database includes the following:

- A user interface through which unsophisticated users communicate with the database.

- An intelligent coupler that serves as an intermediary between the end user and the distributed database system. The coupler listens to the user's retrieval requests; helps the user formulate his requests correctly; efficiently translates user's retrieval requests into a network-compatible retrieval command language; and obtains authorization from the system for data retrieval and/or update.

- Tools for picture data management. Graphical indexing techniques are provided so that the clinical researcher and physician can easily retrieve pictorial/graphical information from the medical database.

- Means for logical database synthesis. This involves conversion of the user's view of the database into a logically coherent physical organization.

C.2.b) Development of a computer-based consultation system for diagnosis and management of glaucoma. This involves ongoing collaboration between Dr. Jacob Wilensky at MISL, and, through SUMEX-AIM, other investigators around the United States. Included are the original investigators in glaucoma consultation: Dr. Casimir Kulikowski (Rutgers), Dr. Shalom Weiss (Mt. Sinai Hospital, NY), and Dr. Aaron Safir (Mt. Sinai Hospital).

C.2.c) Prof. Brian Phillips has used SUMEX-AIM during the last year to implement a model of knowledge, called an encyclopedia. Implemented in INTERLISP, the model includes a network representation for procedures, similar in spirit to the representation in Earl Sacerdoti's "Structure for plans and behavior". Prof. Phillips will discuss his work at an AI conference in Europe early this summer.

C.2.d) Formal models for consultation systems. Petri nets have been studied, primarily be Murata (see below), as a formal representation for interacting parallel processes. Petri nets are similar to causal networks, as described by Kulikowski and Weiss at Rutgers, except that, with Petri nets, cyclic activity is easily represented. The similarity between Petri nets and inference nets has also been noted (Walser and McCormick). The utility of the Petri net framework for modelling physical processes was explored by Walser, with the construction of a simulated coffee maker. Further studies are planned.

D.) LIST OF MISL PUBLICATIONS

Chang S. K., "Intelligent coupling of the user to medical information systems", PROCEEDINGS OF THE THIRD ILLINOIS CONFERENCE ON MEDICAL INFORMATION SYSTEMS, November 4-6, 1976, Chicago, Illinois.

Chang S. K., "A model for distributed computer systems design", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, SMC-6, no. 5, May 1976, pp. 344-359.

Chang S. K. and Cheng W. H., "Database skeleton and its application to logical database synthesis", IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol SE-4, no. 1, 1978, pp. 18-30.

Chang S. K. and Ke J. S., "Database skeleton and its application to fuzzy query translation", IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol SE-4, no. 1, 1978, pp. 31-44.

Chang S. K., Donato N., McCormick B. H., Reuss J. and Rocchetti R., "A relational database system for pictures", PROCEEDINGS IEEE COMPUTER SOCIETY WORKSHOP ON PICTURE DATA DESCRIPTION AND MANAGEMENT, Chicago, Illinois, April 1977.

Church R. W. and Murata T. "Properties of a modified incidence matrix for computation graphs", PROCEEDINGS OF THE 1977 ALLERTON CONFERENCE ON COMMUNICATION, CONTROL, AND COMPUTING, September 1977.

Manacher G. K., "An improved version of the Cocke-Younger-Kasami Algorithm", to be published in J. OF COMPUTER LANGUAGES, 14 pages.

Manacher G. K., "An application of pattern matching to a problem in geometrical complexity", INFORMATION PROCESSING LETTERS, vol. 5, no. 1, 1976, pp. 6-7.

McCormick B. H., Chang S. K., Borovec R. T., Read J. S. and Wilensky J. T., "Technological trends in clinical information systems", MEDINFO 77, Shires and Wolf (Editors), North Holland Publishing Company, 1977, pp. 43-48.

McCormick B. H. and Amendola, R. C., "Cytospectrometers for subcellular particles and macromolecules: design considerations", Workshop on THEORY, DESIGN AND BIOMEDICAL APPLICATIONS OF SOLID STATE CHEMICAL SENSORS, Case Western Reserve University, March 28-30, 1977.

McCormick B. H., Read J. S., Borovec R. T., Amendola R. C., and Goldbaum M. H., "Image processing in Television Ophthalmoscopy", in DIGITAL PROCESSING OF BIOMEDICAL IMAGES, K. Preston and M. Onoe (Editors), Plenum Press, 1976, pp. 399-424.

Murata T., "State equations for E-Net interpreted marked graphs", PROCEEDINGS OF THE 19th MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS, August 1976, pp. 152-157.

Murata T. and Shah T., "on liveness, deadlock, and reachability of E-Nets", PROCEEDINGS OF THE 14th ALLERTON CONFERENCE ON CIRCUIT AND SYSTEM THEORY, October 1976, pp. 597-605.

Murata T., "A method for synthesizing marked graphs from given markings", PROCEEDINGS OF THE 10th ASILOMAR CONFERENCE ON CIRCUITS, SYSTEMS, AND COMPUTERS, November 1976, pp. 202-206.

Murata T., "State equations, controllability and maximal matchings of Petri nets", IEEE TRANSACTIONS ON AUTOMATIC CONTROL, vol. AC-22, No. 3, June 1977, pp. 412-416.

Murata T., "Petri nets, marked graphs, and circuit system theory - a recent CAS application", CIRCUITS AND SYSTEMS, vol. 11, no. 3, June 1977, pp. 2-12.

Murata T., "Circuit theoretic analysis and synthesis of marked graphs", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, vol. CAS-24, no. 7, July 1977.

Murata T. and Onaga K. "Deadlock in capacitied computation networks", PROCEEDINGS OF THE 20th MIDWEST SYMPOSIUM ON CIRCUITS AND SYSTEMS, August 1977.

Murata T. and Onaga K., "Marked graphs with token capacity constraint", IEEE TRANSACTIONS ON CIRCUITS, SYSTEMS, AND COMPUTERS, November 1977.

Murata T., Onaga K., and Tani K., "Maximum firing number in capacited marked graphs", PROCEEDINGS OF THE 1977 ASILOMAR CONFERENCE ON CIRCUITS, SYSTEMS, AND COMPUTERS, November 1977.

Nordmann B. H. and McCormick B. H., "Modular asynchronous control design", IEEE TRANSACTIONS ON COMPUTERS, March 1977.

Tani T., Onaga K. and Murata T., "Termination conditions for capacited computation graphs", PROCEEDINGS OF THE 1977 JAPAN NATIONAL CONFERENCE ON INFORMATION PROCESSING, August 1977.

Walser R. L., Hain T. and McCormick B. H., "Design of a relational database system for pathology", To appear in PROCEEDINGS OF THE FOURTH ILLINOIS CONFERENCE ON MEDICAL INFORMATION SYSTEMS, Urbana, Illinois, May 10-12, 1978.

Walser R. L., McCormick B. H., Chang S. K., Ruppel J., and Wilensky J. T., "Applications of an ophthalmic information system", To appear in PROCEEDINGS OF THE WORKSHOP ON COMPUTER LABORATORY HEALTH CARE RESOURCES PROGRAM, January 11-13th, 1978, Lubbock, Texas.

Walser R. L. and McCormick B. H., "A system for priming a clinical knowledge base", PROCEEDINGS OF THE NATIONAL COMPUTER CONFERENCE, June 1977, Dallas, Texas, pp. 171-177.

Walser R. L. and McCormick B. H., "Organization of clinical knowledge in MEDICO", PROCEEDINGS OF THE THIRD ILLINOIS CONFERENCE ON MEDICAL INFORMATION SYSTEMS, Chicago, Illinois, November 4-6, 1976.

Wilensky J. T., "Intraocular lenses", in OPHTHALMOLOGY: PRINCIPLES AND MANAGEMENT, G. Peyman, D. Sanders, and M. F. Goldberg (Editors), in preparation.

Wilensky J. T., "Current concepts in primary angle-closure glaucoma", ANN. OPHTHALMOL., August 1977, pp. 963-972.

Wilensky J. T., Goldberg M. F. and Alward P., "Glaucoma after pars plana vitrectomy", TRANS. AM. ACAD. OPHTHALMOL. OTOLARYNGOL., vol. 83, Jan.-Feb. 1977, pp. 114-121.

E.) FUNDING STATUS

1. Title: Medical Information Systems Laboratory
2. Principal Investigators: B. H. McCormick and M. F. Goldberg
3. Agency: PHS, Department of Health Manpower
4. Id number: PHS-MB00114

5. & 6. Current terms:
This year (7/77 - 6/78) $222,487
MISL funding ends 6-29-78. MISL will retain its identity however, and much of the work begun under MISL will continue. Many of the individuals who formerly worked under MISL will continue on database projects under ARPA funding.

II) INTERACTION WITH SUMEX-AIM RESOURCE

A.) COLLABORATION

Major collaboration at present is through the ONET, involving the ophthalmology departments of five medical schools. Dr. Jacob Wilensky is actively engaged in evaluating and modifying the Glaucoma Consultation Program, written originally by Shalom Weiss.

B.) CRITIQUE OF RESOURCE SERVICES

Users at MISL are pleased with SUMEX-AIM services. The availability of up-to-date online documentation makes it easy to learn how to use the system and stay abreast of new developments. The online bulletin board is especially commendable. Since documentation is so readily available, consultation with SUMEX staff has rarely been necessary.

III) FOLLOW-ON SUMEX GRANT PERIOD

A.) LONG RANGE USER PROJECTS AND GOALS

One of MISL's accomplishments is the establishment of an information system for the Illinois Eye and Ear Infirmary (IEEI). After MISL funding terminates in June, 1978, the IEEI will continue to use MISL facilities, mainly for clinical databases and text editing in clinical offices. In July (under ARPA funding), several MISL personnel are slated to begin development of an augmented version of RAIN, the relational database system developed under MISL. This new system will form the basis for a relational database in pathology, which will serve both the Illinois Eye and Ear Infirmary and the pathology department at the University of Illinois Hospital.

B.) JUSTIFICATION FOR CONTINUED USE OF SUMEX

Access by our staff to SUMEX facilities and opportunity for inter-institutional collaboration have been enhanced by a SUMEX (PDP-10) - MISL (PDP-11) phone connection, which has been operational since last summer.

One of MISL's original goals was network-compatible database design. We have recently taken a large step toward that goal with the implementation of a special network version of the Unix operating system for our PDP-11 minicomputers. We presently have a network of five minicomputers, including various databases, on the University of Illinois Circle and Medical Center campuses. We have been collecting glaucoma and general eye index data for some time, and will soon create a large pathology database. We feel that we are in an especially good position to experiment with the sharing of data through computer networks. Especially interesting is the sharing of data between very remote sites, using phone lines. Our link to SUMEX-AIM provides a perfect outlet for our distributed database plans.

It is worth noting that several MISL personnel, including Prof. S. K. Chang, R. L. Walser, A. C. Petersen, and H. Dreizen, will continue to work on distributed database problems. The work will continue under the Distributed Image Management And Projection project (DIMAP), funded by ARPA. While the former MISL personnel will take on new obligations, they will most certainly continue to interact with physicians at the Illinois Eye and Ear Infirmary. It is planned that much of the DIMAP research will take place on the former MISL computer system; the DIMAP group has agreed to maintain the Unix operating system in exchange for free access to the computer system.

SUMEX-AIM is presently being used extensively by Prof. Brian Phillips and his students. Using INTERLISP, Prof. Phillips is coding a model of knowledge developed over a period of years at the State University of New York at Buffalo, and later in the Department of Information Engineering and MISL in Chicago. While the model of knowledge is well-developed, and has been implemented at another site in SNOBOL, the INTERLISP version requires further work. Since Prof. Phillips has shifted his work entirely to SUMEX-AIM, he is quite anxious to keep his access privileges.

C.) SUGGESTIONS FOR FUTURE RESOURCE DEVELOPMENT EFFORTS

As mentioned above, we are very interested in coupling our PDP-11 based UNIX operating system with the SUMEX-AIM network. and would like to encourage similar connections at other sites. There are several advantages. Maintaining voluminous patient-related data on minicomputer systems would provide for local security, and help to keep SUMEX secondary storage free for service and development programs and documentation. The enhanced opportunity for inter-site collaboration and database sharing is obvious, and would be beneficial to the SUMEX-AIM community as a whole.

4.1.6 PUFF/VM PROJECT

PUFF/VM: Pulmonary Function and Ventilator Management Project

John J. Osborn, M.D.
The Institutes of Medical Sciences (San Francisco)
Pacific Medical Center

and

Edward A. Feigenbaum
Computer Science Department
Stanford University

The immediate goal of this project is the development of knowledge-based programs which interpret physiological measurements made in clinical medicine. The interpretations are intended to be used to aid in diagnostic decision making and in therapeutic actions. The programs will operate within medical domains which have well developed measurement technologies and reasonably well understood procedures for interpretation of measured results. The programs are:

(1) PUFF: the interpretation of standard pulmonary function laboratory data which include measured flows, lung volumes, pulmonary diffusion capacity and pulmonary mechanics, and

(2) VM: management of respiratory insufficiency in the intensive care unit.

The second, but equally important, goal of this project is the dissemination of Artificial Intelligence techniques and methodologies to medical communities that are involved in computer aided medical diagnosis and interpretation of patient data.

I. SUMMARY OF RESEARCH PROGRAM

PUFF

A. Technical Goals

The task of PUFF program is to interpret standard measures of pulmonary function. It is intended that PUFF produce a report for the patient record, explaining the clinical significance of measured test results. PUFF also must provide a diagnosis of the presence and severity of pulmonary disease in terms of measured data, referral diagnosis, and patient characteristics. The program must operate effectively over a wide range of pathological conditions with a broad clinical perspective about the possible complexity of the pathology.

B. Medical Relevance and Collaboration

Interpretation of standard pulmonary function tests involves attempting to identify the presence of obstructive airways disease (OAD: indicated by reduced flow rates during forced exhalation), restrictive lung disease (RLD: indicated by reduced lung volumes), and alveolar-capillary diffusion defect (DD: indicated by reduced diffusivity of inhaled CO into the blood). Obstruction and restriction may exist concurrently, and the presence of one mediates the severity of the other. Obstruction of several types can exist. In the laboratory at the Pacific Medical Center (PMC), about 50 parameters are calculated from measurement of lung volumes, flow rates, and diffusion capacity. In addition to these measurements, the physician may also consider patient history and referral diagnosis in interpreting the test results and diagnosing the presence and severity of pulmonary disease.

Currently PUFF contains a set of about 60 physiologically based interpretation "rules". Each rule is of the form "IF <condition> THEN <conclusion>". Each rule relates physiological measurements or states to a conclusion about the physiological significance of the measurement or state.

The interpretation system operates in a batch mode, accepting input data and printing a report for each patient. The report includes: (1) Interpretation of the physiological meaning of the test results, the limitation on the interpretation because of bad or missing data; the response to bronchodilators if used; and the consistency of the findings and referral diagnosis. (2) clinical findings, including the applicability of the use of bronchodilators, the consistency of multiple indications for airway obstruction, the relation between test results, patient characteristics and referral diagnosis. (3) Interpretation Summary, which consists of the diagnosis of presence and severity of abnormality of pulmonary function.

Initially the rules were proposed by our pulmonary physiologist during discussions with systems staff consisting of biomedical engineers from PMC and computer scientists from Stanford. The process of rule development involved iteratively refining the rules to make the rule interpretations arbitrarily close to the interpretation by the physiologist for a set of test cases. A set of 107 representative case results was selected from the files of our laboratory for retrospective analysis during rule development. Statistical analysis was used to characterize the consistency and information content of the retrospective data. Prospective analysis of 144 new cases then was used to evaluate the consistency of the interpretations by the physiologist in comparison with a second independent physiologist , and with the rules.

C. Progress Summary

Knowledge base:

PUFF is implemented on the PDP10 in a version of the MYCIN system which is designed to accept rules from new task domains. Currently approximately 60 pulmonary physiology rules related to the interpretation of measurements mentioned above have been implemented. A typical rule is:

```
If (FVC(PP))>=80) and  (FEV1/FVC<predicted-5) then PEAK FLOW RATES ARE
REDUCED, SUGGESTING AIRWAY OBSTRUCTION OF DEGREE
        if (predicted-15<= FEV1/FVC <predicted-5) MILD
        if (predicted-25<=FEV1/FVC <predicted-15) MODERATE
        if (predicted-35<=FEV1/FVC <predicted-25) MODERATE TO SEVERE
        if (FEV1/FVC <predicted-35)               SEVERE
```

Results

A representative sample of 107 test cases was taken from the computer records of our pulmonary function laboratory. During the process of characterizing the data base statistically and developing the rule set for interpretation, the director of our pulmonary laboratory changed his diagnosis on 42 of the 107 cases. The largest number (14) of the changes in diagnosis was a change from an original diagnosis of Pulmonary Function "Within Wide Limits of Normal" to a diagnosis of obstruction or restriction of a mild degree. The original normal diagnosis was made either because he felt that the patient did not need to be confronted with an abnormal diagnosis, or because he knew the patient and made the diagnosis on background information rather than the measured data. The addition of diffusion defects or changes in degree were involved for nine patients. In five cases a mild restriction diagnosis was added for patients with OAD. Three data errors were found, and three OAD subtype changes were made. The majority of the remainder (14) of the changed diagnoses were changes of degree of severity by one degree, e.g. from mild to moderate. The changes in severity were made because, when he analyzed the data systematically, he came to a different conclusion than when individual cases were analyzed in isolation.

The results of a preliminary comparison we conducted to compare the interpretations made by PUFF and those made by the physiologists are summarized in Tables 1 and 2. Table 1 compares agreement in diagnoses made by the first MD-pulmonary physiologist with the second physiologist and with the PUFF rules. Comparison is made independent of the severity of condition reported in the diagnosis. The overall rate of agreement between the two physiologists on four diagnoses (Normal, OAD, RLD, DD) for 144 cases was 86%, and between the first physiologist and the rules 90%. There was substantially lower percentage of agreement in diagnosing OAD subtype (asthma, bronchitis, emphysema) than in diagnosing primary pulmonary abnormality.

Diagnosis (Dx)	Number: With Dx	Number: Without Dx	Percent Agreement With Diagnosis by 1st M.D.: Second M.D.	Percent Agreement With Diagnosis by 1st M.D.: PUFF Rules
Normal	31	113	0.91	0.93
OAD	79	65	0.86	0.88
RLD	52	92	0.92	0.94
DD	40	104	0.76	0.86
Asthma	15	64	0.81	0.84
Bronchitis	9	70	0.61	0.81
Emphysema	32	47	0.82	0.84

Table 1. Percentage agreement of diagnoses by a second MD-physiologist and by PUFF rules with a first MD-pulmonary physiologist. 144 cases were interpreted prospectively. Some abnormal cases had more than one pathological condition.

Table 2 compares agreement in severity of diagnoses made by the two MD's and by PUFF rules. In 94% of the 144 cases, the degree of severity (0=none; 1=mild; 2=moderate; 3=moderately-severe; 4=severe) of OAD diagnosed by the first MD was within a single degree of severity of OAD diagnosed by the second MD. In 96% of the 79 cases for which the first MD diagnosed OAD, the second MD diagnosed the severity of OAD within one level of the severity diagnosed by the first MD. Agreement within one degree of severity of the diagnoses by the first and second MD's was substantially lower in RLD and DD cases. These discrepancies occurred because the second MD consistently called RLD more severe than did the first MD, and he consistently did not diagnose diffusion defects when the first MD diagnosed DD of moderate or greater degree.

	Percent Agreement with 1st MD			
	All 144 cases		1st MD made Dx	
Diagnosis	Second M.D.	PUFF Rules	Second M.D.	PUFF Rules
Normal				
OAD	0.94	0.99	0.96	0.97
RLD	0.92	0.97	0.77	1.00
DD	0.87	0.87	0.60	0.80
Total	0.91	0.94	0.86	0.94

Table 2. Percent agreement within one degree of severity of diagnoses by two MD's and by the first MD and rules.

Tables 1 & 2 are taken from analysis of the same 144 cases.

VM

A. Technical Goals

The task of the VM program is to provide real-time consultative advice for patients undergoing mechanical ventilation in the intensive care unit at the Pacific Medical Center. This intensive care unit has an on-line computer based patient monitoring system which automatically obtains approximately twenty physiological parameters. VM is intended to be an extension of this system that will (1) provide a summary of the patient status easily understood by the clinician, (2) recognize untoward events and provide suggestions for corrective action, (3) advise on adjustment of a mechanical ventilator based on assessments of patient status and therapeutic goals, and (4) detect possible measurement errors.

The VM system is designed to be an expectation-driven system which utilizes the current and past history of the patient to establish guidelines for interpreting patient measurements. These guidelines, or expectations, are used to dynamically establish upper and lower limits for comparison with each measurement. Persistent patterns in the violations of these limits represent unexpected clinical situations that may require adjustment of the equipment or other actions by the physician..

The overall plan includes a telephone connection between the IBM 1800/PDP-11 patient monitoring system at PMC and the Sumex PDP-10. The ICU patient's physiological measurements will be provided on a 2-10 minute basis for analysis by the VM program. Summary information, suggestions to the clinicians, and/or requests for additional information will be sent back to the ICU for action.

B. Medical Relevance and Collaboration

The VM problem domain of interpretation in the dynamic ICU environment brings together the designers who created the current monitoring system with computer scientists experienced in medical interpretation (MYCIN) and signal understanding research [Shortliffe:76, Nii:77]. We are directing our attention to the assistance of clinicians in the management of patients who need ventilatory support. VM will offer advice on a wide range of clinical problems ranging from one-time decisions on the selection of an appropriate ventilator and the initial settings for the equipment to continuous monitoring of the patient. The main effort is directed toward the tracking of the patient physiological status while on the ventilator, and on providing suggestions for adjustments of the ventilator settings as necessary. This includes advice on the process of removing the patient from the ventilator. An up-to-date summary stated in terms of physiologically based conclusions rather than the individual measurements will also be provided.

To assist in the interpretation process, VM must be able to recognize unusual or unexpected clinical events (including machine malfunction) in a manner specifically tailored to the patient in question. The interpretation task is viewed as an ongoing process in the ICU, so that the physiological measurements must be continually reevaluated producing a current clinical picture.

This picture can then be compared with previous summary of patient status to recognize changes in patient condition upon which therapy selection and modifications can be made. The program must also determine when the measurements are most likely to be sensitive to error or when external measurements would be of diagnostic significance.

VM offers a new approach towards more accurate recognition of alarm conditions by utilizing the history and situation of the patient in the analysis. This is in contrast to the use of static limits applied to measurements generated to fit the "typical patient" under normal conditions. Our program uses a model of interpretation process, including the types and levels of conclusions drawn manually from the measurements to provide a summary of patient condition and trends. The program generated conclusions are stated at levels more abstract than the raw data; for example, the presence of hemodynamic stability/instability rather than in terms of heart rate and mean arterial pressure. When the data is not reliable enough to make these conclusions, additional test may be suggested. The recognition of important conclusion for which external verification is sought, will also elicit the suggestion for confirming tests from the program.

C. Progress Summary

A prototype system is currently running using real data from the ICU provided on magnetic tape. The program uses a small set of rules (currently around 30) to suggest appropriate time to proceed in the weaning process (i.e. weaning the patient from the ventilator's assistance). It also provides summaries and suggestions about the patient status and recent trends.

A large amount of the work has gone into determining (1) an appropriate representation of knowledge that can deal with the dynamic nature of this problem, and (2) the extent of knowledge needed to develop a satisfactory level of performance in the program. Because this program must run in a near real time environment, a large segment of the design is concerned with rules which can focus quickly on knowledge and processes which are relevant for different patient situations.

E. Funding

Approved by NIH but not yet funded.

II. <u>RESEARCH PLANS</u>

A. Long Range goals and plans

<u>PUFF</u>

Performance Improvement

With the existing PUFF rules, the pulmonary physiologist at PMC is able to process approximately 50% of the patients undergoing the pulmonary function tests. Some of the system interpretations are used with relatively minor stylistic changes in the printed interpretations, while other cases require important points to be added to the system interpretation. We would like to increase the utility of of PUFF to where it can process 90-95% of the patients. In order to reach this goal the current program must be augmented with rules about:

(1) restrictive lung disease - the current rules do not identify RLD with sufficient accuracy.

(2) modify some of the existing rules on OAD,

(3) add rules to determine patient effort, or lack of effort, during the measurement acquisition,

(4) add rules related to blood gas analysis, and

(5) modify some parts of the PUFF program to increase the efficiency.

Consensus

Physician acceptance of assistance by knowledge-based programs is understandably inhibited by disagreements between physician diagnoses and those produced by the programs. This disagreement reflects a deeper underlying disagreement among the physicians themselves on the rules to be used for diagnosis. A more subtle problem arises when physicians agree on the diagnosis but cannot agree on the supporting evidence or the reasoning which led to the diagnosis. Currently, two pulmonary physiologists are attempting to obtain consensus between themselves and with the current rules in PUFF.

We consider obtaining consensus among physicians an important problem for PUFF's acceptance. At the same time, the process of building consensus about a body of knowledge is an interesting area of research in artificial intelligence. The general question to be answered is: What components of the diagnostic process cause differences in the diagnosis? More specifically, in terms of the knowledge which the physicians bring to bear on the diagnostic process, are there differences in representation? (i.e. do physicians use different forms of knowledge?) Are there differences in the definition? (i.e. do physicians use different models or definitions of disease, manifestations, and diagnosis?) Are there differences in the process? (i.e. do physicians reason differently about the problem?)

A second set of questions are: When designing knowledge-based systems, should it be designed to be modifiable to conform to the user physician's diagnostic process? Or, should it be designed to act to identify and record areas of disagreement among the physicians? Should it warn the physicians, mediate the differences, or merely collect the information?

We would like to spend the next few years answering some of these questions in order to understand the problem of consensus building among physicians.

VM

The long range goal of the VM project is to develop and evaluate an extensive rule base sufficient to improve patient care in the ICU and augment the effectiveness of the current ICU monitoring system. A large amount of effort will be expended in creating a knowledge base capable of handling the various unexpected events which may occur in the Intensive Care Unit to critically ill patients. The knowledge base must also encompass decisions for selection of ventilatory assistance equipment and other important clinical decisions connected with recognition of the patient situation and prognosis.

To insure acceptance by physicians, a careful evaluation of the correctness of the advice of the program against a set of prospective cases will be carried out. A large amount of patient data have already been collected for this evaluation. An experimental period of use in the ICU by clinicians associated with our project is projected before the end of 1980.

The VM program embodies new techniques for knowledge-based programs capable of handling problems dealing with series of interpretations under real time output requirements. The program design should also provide a basis for other medical programs which can handle patient data which change over time, not necessarily in the context of the ICU. We intend the design of the VM program to be independent of the knowledge utilized. The characterization of similar domains to which this new technology can be applied will be an on-going goal of this research team.

III. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations and medical use of programs via SUMEX

The PUFF/VM project requires very close collaboration between investigators at two institutions separated by fifty miles. This kind of collaboration, in which program development and testing proceeds concurrently on the same application system, requires a network facility for sharing of code, data and ideas. SUMEX has been used at PMC for running programs developed concurrently by Stanford and PMC staff, and data has been taken from the PMC computer system and transferred to SUMEX on magnetic tape for program development and testing.

B. Sharing and interactions with other SUMEX-AIM projects

We have participated in the AIM workshop and had very fruitful interaction with a number of other SUMEX users, directly influencing our perception of important problems and potentially appropriate solutions. Personal contacts at other conferences, at Stanford AI weekly meetings, and at PMC with visiting members of the AIM community, have also been very helpful in keeping abreast of the current thinking of other members of the AI community and with members of the medical community interested in computer based physiological analysis and diagnosis. Specifically, there is the closest possible collaboration with researchers on the MYCIN, MOLGEN and DENDRAL projects, who share common space, common techniques, and common attitudes.

C. Critique of resource management

The SUMEX community continues to be an extremely supportive environment in which to do research on uses of artificial intelligence in clinical medicine. The community has two equally vital resources -- the people with knowledge and interest in AI and the facility on which AI system development can proceed. They are equally excellent as resources, helping hands when faced with problems, and friendly support for continued productive research. The availability of INTERLISP; of a facility on which routine data processing functions (eg. manipulating magnetic tapes and making long listings) can take place; and of message-sending among remote users are all vital functions for our project. SUMEX provides them in an environment which is friendly and reliable.

D. Needs and plans for other computational resources

The computation facility at PMC is currently the source of all of the data being used by the PUFF/VM project, and it will continue in this capacity. We expect to link the two machines using a simple telephone dial-up link, but this represents the only system increment to the computational facility of the collaborative project. As the AI techniques developed under PUFF/VM enter routine clinical use at PMC, we have the requirement for system support on which these programs can execute. To date, we have been able to use the existing software development facilities on the PMC PDP-11 for this purpose, but we hope to be able to use more powerful mini-computer based software development facilities in the future.

E. Recommendations for future community and resource development

We perceive the evolution of our AI capability as moving from a highly speculative development state, for which the interactive development capabilities of SUMEX are vital, to a more stable but still changing validation-and-evaluation state. Ultimately we foresee rather stable specification of a program for routine clinical use. Thus, we see the need to transfer our AI techniques from the SUMEX PDP-10 to a local host. For this transfer, a principal long-range need is for software systems that will allow us to run AI systems on a mini-computer after they have been developed on the more powerful SUMEX facility. If the validation of PUFF/VM in the PMC clinical setting shows the programs to be effective in health care, then we hope and expect to be able to provide the capability on a routine basis.

4.1.7 RUTGERS COMPUTERS IN BIOMEDICINE

Rutgers Research Resource - Computers in Biomedicine

Principal Investigator: Saul Amarel
Rutgers University, New Brunswick, New Jersey

I. SUMMARY OF RESEARCH PROGRAM

A) Goals and Approach

The fundamental objective of the Rutgers Resource is to develop a computer based framework for significant research in the biomedical sciences and for the application of research results to the solution of important problems in health care. The focal concept is to introduce advanced methods of computer science - particularly in artificial intelligence - into specific areas of biomedical inquiry. The computer is used as an integral part of the inquiry process, both for the development and organization of knowledge in a domain and for its utilization in problem solving and in processes of experimentation and theory formation.

The Resource community includes 51 researchers - 30 members, 7 associates, and 14 collaborators. Members are mainly located at Rutgers. Collaborators are located in several distant sites and they interact, via the SUMEX-AIM and RUTGERS-10 facilities, with Resource members on a variety of projects, ranging from system design/improvement to clinical data gathering and system testing. At present, collaborators are located at the Mt. Sinai School of Medicine, N.Y.; Washington University School of Medicine, St. Louis, Missouri; Johns Hopkins Medical Center, Baltimore, Maryland; Illinois Eye and Ear Infirmary, Chicago, Illinois; the University of Miami, Florida, and the University of Pennsylvania, Philadelphia, Pennsylvania.

Resource activities include research projects (collaborative research and core research), training/dissemination projects, and computing in support of pilot user projects. The research projects are organized in three main AREAS OF STUDY. The areas of study and the senior investigators in each of these are:

(1) Medical Modeling and Decision Making (C. Kulikowski).

(2) Modeling Belief Systems and Commonsense Reasoning (C. F. Schmidt and N. S. Sridharan).

(3) Artificial Intelligence: Representations, Reasoning, and System Development (S. Amarel).

The training/dissemination activities of the Rutgers Resource (coordinated by R. Smith) include sponsorship of the Annual AIM Workshop - whose main objective is to strengthen interactions between AIM investigators, to disseminate research methodologies and results, and to stimulate collaborations and imaginative resource sharing within the framework of AIM. The third AIM Workshop was held at Rutgers on July 5 to 8, 1977. The fourth AIM Workshop is scheduled for June 25 to 28, 1978.

The RUTGERS-10 computer is being used not only for support of local research projects and AIM Workshop activities; but also for Pilot User Projects in the AIM community - within the general framework of the national AIM project. Computing activities in the Resource are coordinated by S. Levy.

B) Medical Relevance; Collaborations

A unique and novel aspect of our work is the creation of a network of clinical investigators to collaborate on the testing and continued development of the computer programs needed to accomplish the tasks mentioned in Section C.1.a. During 1977, the ophthalmological network (ONET) of glaucoma investigators continued to grow and has established itself, with several significant collaborative research projects currently under way. The consultation program for glaucoma using the causal association network (CASNET) model developed within the Rutgers Resource, was jointly presented by the ONET members at the 1976 meeting of the Association for Research in Vision and Ophthalmology. The results of the panel discussion where the CASNET/Glaucoma program was pitted against a group of experts are described in the recently published book Discussions On Glaucoma by Lichter and Anderson. The incorporation into the consultation program of alternative expert opinions on subjects currently under debate has been an important aspect of this work. The SUMEX-AIM shared computer resource has been essential to the activities of ONET.

The knowledge base and the strategies of our CASNET glaucoma consultation system are being strengthened and refined continuously in the ONET environment. The system is now at a point where it is considered by leading ophthalmologists as "highly competent-to-expert" in several subspecialties of glaucoma.

In the past year work has begun in several new research areas: hematology, endocrinology and enzyme kinetics. In addition to the design of consultation systems, underlying problems of knowledge representation and inference processes have been studied.

Research collaborations include the five medical centers of ONET, the Rutgers Medical School, and the University of Pennsylvania.

In the area of Belief Systems, collaboration is continuing with Prof. Andrea Sedlak and her research group at the University of North Carolina at Chapel Hill. Also, interactions continued with a group at Univ. of Massachusetts on the design of a knowledge-based system to learn rules for interpreting actions.

Our close contacts with the Stanford projects on Heuristic Programming (Drs. Buchanan, Feigenbaum, Lederberg) are continuing. The AI orientation and approach of these Stanford projects are very similar to ours. Graduate students at Rutgers continued to contribute to the development of the AI handbook - a project led by Dr. Feigenbaum at Stanford and intended to provide a network-accessible encyclopedic coverage of the AI field for the AIM community and AIM guests.

C) Progress Summary

1. Areas of Study and Projects:

a) Medical Modeling and Decision-Making

Research activities during the past year have concentrated on the development of new consultation systems, and associated investigations into representations of knowledge, strategies of inference, and planning. The evaluation and testing of the CASNET/Glaucoma system by the members of the ONET (Ophthalmological Network) has continued. Several new application areas for consultation have been started: hematology, endocrinology, and most recently, rheumatology. The methodology of modeling in enzyme kinetics has been studied, and preliminary data base and research support programs developed.

The collaborative activities of ONET, comprising investigators at the Mt. Sinai School of Medicine, Johns Hopkins University Washington University, the University of Illinois at Chicago, and the University of Miami, have continued through the testing of the CASNET/Glaucoma Model. The time-sequenced data base facilities have been used both by ONET and by hemophilia investigators at the College of Medicine and Dentistry of New Jersey (Rutgers Medical School). Investigators at this institution have also collaborated on the development of disease models for consultation in endocrinology. Collaborative investigation at the University of Pennsylvania has begun on problems of enzyme kinetics modeling.

In the area of general methods and systems for medical reasoning, there has been research into the representation of anatomical physiological knowledge in the form of descriptive models that interact with clinical-level models for decision-making. The AIMDS descriptive system has been used to model the visual pathways and to implement the generalized physiological reasoning that permits the interpretation of patterns of visual field loss. A related project involves the 3-dimensional graphical representation of the 'hill of vision' for clinical interpretation.

The investigation of reasoning schemes for diagnosis and treatment has been carried out (and new representational elements included) in two systems currently under development: IRIS and XPERT. The former is a semantic network-based system for propagating inferences written in INTERLISP. The latter is a modular consultation system that is being designed for ease in knowledge acquisition from experts and for transferability to small machines (the PDP-11/60 is being used as a prototype).

Clinical investigations in hemophilia, thyroid disease and hypertension have been aided by Resource support and development of the BRIGHT system.

b) Modeling of Belief Systems and Commonsense Reasoning

The last year has been a period of intensive system development in AIMDS and one of intensive conceptual development in BELIEVER.

AIMDS is intended to be a representational framework, which provides a combination of description languages, a logical language and several major procedures for creating and matching large description structures, and facilitates the construction of a variety of knowledge-based systems. Heavy emphasis is given to the representation and modeling of actions which may have side effects. A context mechanism is available that makes it possible to store and process plans, maintain belief models, and answer "what if" questions about actions. The intent is to have the user describe the knowledge being applied as needed, and consequently are expected to be simple to build, test, and also to understand.

The BELIEVER system is an example of a psychological theory that uses the tools of artificial intelligence, to construct a psychological theory of the process of understanding observed actions of others. Understanding observed action sequences involves the attribution of a plan to the actor and identifying the actor's goals. A plan is attributed to the actor even if the plan is not executed to completion. The attributed plan must be a plan that would realize the given goal, with respect to the observer's model of the actor's beliefs.

The plan must correspond to the observed actions. The goal identified must be attributable to the actor, i.e., there must be a motivational reason derivable for the goal for the actor. The evidence from subject experiments of recall and summarization suggest strongly that the observers develop a model of the actor's beliefs and intentions and that they focus on one hypothesis about the actor's plan. If the subsequent actions match the expectations this serves to confirm parts of the plan and allows the observer to refine the plans used for monitoring the action sequence. If subsequent actions disconfirm the expectations, this requires revisions to the hypothesized plan structure (perhaps including the identified goal), attempting to retain as much of the original interpretation as is possible.

The principal activity in BELIEVER has been the further clarification given to the notion of the "Hypothesize and Revise" paradigm and its articulation for the Plan Recognition problem in the form of about a dozen hypothesis revision rules. The arguments for the paradigm and a description of the revision rules are given in a paper that is to appear.

The BELIEVER and AIMDS system are programmed in a combination of RUCILISP and FUZZY. All of the effort involved in the design, construction, and testing of the psychological theory are carried out on the RUTGERS-10.

c) Artificial Intelligence; Representations, Reasoning, and Systems Development

Our work in this area continues to be oriented to collaboration with investigators in other Resource projects and to study of basic AI problems that are related to Resource applications. The collaborations involve adaptation and augmentation of existing AI methods and techniques to handle specific key problems identified in the application projects.

The close collaboration with investigators in the Belief Systems area has resulted in further development of the AIMDS system for handling problems of action interpretation of the type encountered in the domain of the BELIEVER theory.

The AIMDS system was operational in the summer of 1976. A vastly revised and speeded-up version was running in 1977 The system development of AIMDS in the past year has been directed at (a) the production of a user manual; (b) the production of several tutorial sessions; (c) general user engineering - adding a number of user controlled options, easy specification of these options, and making several of the functions provide helpful text; (d) adding more power in the various description languages that the system provides the user; (e) extending the system to handle multiple contexts in the data base; (f) adding a feature whereby defaults can be computed for missing relations; (g) construction of a more powerful Match function for comparing the description of one object with another. The AIMDS system continues to be an invaluable component of the the research in the construction of the BELIEVER theory. The combination of description and logical languages provided by the system and the various major processes play a significant role in the way the theory is made explicit.

In an attempt to test the generality of the representation available in AIMDS, we have programmed a learning algorithm which learns structural models from examination of examples and near misses, in the way of Pat Winston. The learning program has certain novel features, which eliminates the need to store and process past training instances. We have introduced a separate memory model in which a summary description of the past training instances are maintained. The memory model used for the blocks world is adequate to duplicate the learning phenomena demonstrated by Winston, while operating with a reduced memory load. The memory model is flexible and may be redesigned for different applications quite easily by changing its specifications written in the description language provided by AIMDS. An effort is being made to investigate this form of learning in the domain of C13 NMR spectra. It is anticipated that this form of learning which involves concept refinement by examining examples and near-misses will become another component of the Believer process.

One project currently underway in the general area of natural language interfaces is concerned with the development of a method for the parsing and interpretation of medical notes, and for guiding the researcher in the use and extension of the system's capabilities. By enabling the system to generate and track a closely constrained set of syntactic and semantic expectations during the parsing process, we hope to be able to permit it to deal intelligently with some of the problems arising from syntactic and semantic ambiguity and the widespread ellipsis of medical notes. New schemes for organization of this tracking strategy are being explored.

This year, in a continuing project in hypothesis formation, we have made further progress in the computer acquisition of domain knowledge from empirical data. The domain knowledge is in the form of weighted production rules, and a set of production rules can be represented as a stochastic graph. We obtained a general result about loop-free interpretation of a stochastic graph under max and min operations. We also wrote three programs (one in SITBOL, the other two in FORTRAN) as a preliminary implementation of a system for the semi-automatic construction of a succinct graph model from a large data base, and for the use of a model in knowledge-directed evidence gathering and decision making once it has been constructed. We began the testing of this system of three programs on a data base of case histories of glaucoma patients.

In the area of <u>theory formation in programming</u> our main effort has been directed to representations of programs in various stages of specification, and to methods for acquiring knowledge and for managing a knowledge base for the theory formation system.

During this period work continued on the development of a supportive local programming environment for our research. The FUZZY AI language is now relatively stable. In addition to its use in the Resource, as an implementation tool and as a means of modeling approximate algorithms, FUZZY was exported to a number of other sites in the United States and Western Europe. The RUTGERS/UCI LISP system, on which FUZZY is built, has been improved considerably over the past year. A number of new I/O facilities have been added, including a random access capability, support for sub-file directories, lower-to-upper case character conversion, and user-definable ersatz device names. String storage utilization was improved, and several new string-handling functions were incorporated into the system. The program commenting facility was improved and made storage efficient, and several new source-file handling features were added. A number of new functions and macros were added, including a sort/merge package and an ALGOL-like DO macro. The use of complex macros in interpreted code was optimized via the automatic saving of macro expansions. In addition, the compiler was extensively debugged and its generated code optimized. RUTGERS/UCI LISP is now in use at approximately 25 sites in the United States and Western Europe, and continues to spread. Some of the more active sites (Carnegie-Mellon University, University of California at Irvine, and University of Texas) have actively collaborated with Rutgers in system development activities.

2. AIM Dissemination/Training; The AIM Workshop

The third annual AIM Workshop took place July 5 to 8, 1977, at the Continuing Education Center of Rutgers University, and about 50 invitees attended.

The program included review of recent developments at centers where AI research is being done, with particular emphasis to work in medicine, biochemistry, medical information systems, plus related areas such as vision, speech understanding, and AI in education. In addition, an open meeting of the AIM Advisory Committee took place, and several panel discussions and working groups discussed important technical and management issues in depth. Scheduled panels had as topics AIM Research Management and Knowledge Engineering. As in the past, computer facilities were available for hands-on experimentation and demonstration of AI systems running at various sites such as SUMEX-AIM, SRI, and Rutgers, accessed via the TYMNET and ARPANET.

This year, the Workshop dealt more intensely with a smaller set of issues put to a smaller group of workers in the field. Technical reports were given by those individuals currently most active in AI; the panel discussions were planned to center around research policy and management issues. This format of the Workshop was chosen since it is the view of many senior researchers and representatives of funding groups that AI itself has made much progress, that reasonable formalisms and practical programming tools for knowledge representation and inference are now available, that much experience has been gained in developing systems in Research and Development environments, and that further progress in the direction of "exporting" systems to practical user environments requires careful thought about management and engineering issues.

A feature of the Workshop that was viewed as being particularly successful was the dynamic scheduling of the working groups. Under the direction of Dr. C. Kulikowski, technical program chairman of the Workshop, five working groups met and discussed issues related to knowledge representation and communication and dissemination in more detail; the results of these working groups were reported to the entire workshop on the last day. This format stimulated a good deal of discussion and interaction.

Most of the sessions of the Workshop were videotaped by the Rutgers Instructional Television campus facility. These videotapes are available for viewing, and form a permanent record of the Workshop itself. In addition, the proceedings of the workshop has been prepared, as a summary of the Workshop, and is being distributed to the invitees.

A panel on Applications of AI was organized by S. Amarel from the IJCAI-77 Conference which was held in Boston, Mass. on August 22 to 26, 1977. This panel was intended to further supplement the dissemination activities of AIM by bringing to a wide audience of AI researchers a status report on current development in AI systems and by focusing on some of the key AI problems that were identified so far in AIM research.

3. Computing facilities; Pilot Projects

During this period several new Pilot Projects for the AIM community were initiated on the RUTGERS-10 computer. These include collaboration with SUMEX-AIM on the design and testing of MAINSAIL on our TOPS-10 system, and work on BRIGHT, a NIH-sponsored clinical data base project.

D) Up-To-Date List of Publications

Amarel, S. and C. Kulikowski (1972) "Medical Decision Making and Computer Modeling," Proc. of 5th International Conference on Systems Science, Honolulu, January 1972.

Amarel, S. (1974) "Inference of Programs from Sample Computations," Proc. of NATO Advanced Study Institute on Computer Oriented Learning Processes, 1974, Bonas, France.

Amarel, S. (1974) "Computer-Based Modeling and Interpretation in Medicine and Psychology: The Rutgers Research Resource," Proc. on Conference on the Computer as a Research Tool in the Life Sciences, June 1974, Aspen, by FASEB; also appears as Computers in Biomedicine TR-29. June 1974, Rutgers University, also in Computers in Life Sciences. W. Siler and D. Lindberg (eds.), Faseb and Plenum, 1975.

Amarel, S. (1976) Abstract of Panel on "AI Applications in Science and Medicine" in 1976 National Computer Conference Program, N.Y., June 7-10, 1976.

Amarel, S. (1976) "Summary of National Computer Conference panel on AI Applications in Science and Medicine" in SIGART Newsletter and SIGBIO Newsletter, Fall 1976.

Amarel, S. (1977) "Summary of Issues and Open Problems in AI Applications," in Report of panel on Applications of Artificial Intelligence, Amarel (ed.), in Proc. IJCAI-77, MIT, Cambridge, Aug. 1977; S. Amarel organized this panel session for IJCAI-77.

Bruce, B. (1972) "A Model for Temporal Reference and its Application in a Question Answering Program," in "Artificial Intelligence," Vol. 3, Spring 1972.

Bruce, B. (1973) "A Logic for Unknown Outcomes," Notre Dame Journal of Formal Logic; also appears as Computers in Biomedicine, TM-35, Nov. 1973, Rutgers University.

Bruce, B. (1973) "Case Structure Systems," Proc. 3rd International Joint Conference on Artificial Intelligence (IFCAI), August 1973.

Bruce, B. (1975) "Belief Systems and Language Understanding," Current Trends in the Language Sciences, Sedelow, and Sedelow (eds.) Houston, in press.

Chokhani, S. and C.A. Kulikowski (1973) "Process Control Model for the Regulation of Intraocular Pressure and Glaucoma," Proc. IEEE Systems, Man and Cybernetics Conf., Boston, November 1973.

Chokhani, S. (1975) "On the Interpretation of Biomathematical Models Within a Class of Decision-Making Procedures," Ph.D. Thesis, Rutgers University; also Computers in Biomedicine TR-43, May 1973.

Fabens, W. (1972) "PEDAGLOT. A Teaching Learning System for Programming Language," Proc. ACM Sigplan Symposium on Pedagogic Languages, January 1972.

Fabens, W. (1975) "PEDAGLOT and Understanding Natural Language Processing." Proc. of the 13th Annual Meeting of the Assoc. of Computational Linguistics, October 30 - November 1, 1975.

Kulikowski, C.A. and S. Weiss (1972) "Strategies for Data Base Utilization in Sequential Pattern Recognition," Proc. IEEE Conf. on Decision and Control, Symp. on Adaptive Processes, December 1972.

Kulikowski, C.A. and S. Weiss (1973) "An Interactive Facility for the Inferential Modeling of Disease," Proc. 7th Annual Princeton Conf. on Information Sciences and Systems, March 1973.

Kulikowski, C.A. (1973) "Theory Formation in Medicine: A Network Structure for Inference," Proc. International Conference on Systems Science, January 1973.

Kulikowski, C.A., S. Weiss and A. Safir (1973) "Glaucoma Diagnosis and Therapy by Computer," Proc. Annual Meeting of the Assoc. for Research in Vision and Ophthalmology, May 1973.

Kulikowski, C.A. (1973) "Medical Decision-Making and the Modeling of Disease," Proc. First International Conf. on Pattern Recognition, October 1973.

Kulikowski, C.A. (1974) "Computer-Based Medical Consultation - A Representation of Treatment Strategies," Proc. Hawaii International Conf. on Systems Science, Jan. 1974.

Kulikowski, C.A. (1974) "A System for Computer-Based Medical Consultation," Proc. National Computer Conference, Chicago, May 1974.

Kulikowski, C.A. and A. Safir (1975) "Computer-Based Systems Vision Care," Proceedings IEEE Intercon, April 1975.

Kulikowski, C.A. and Trigoboff, M. (1975) "A Multiple Hypothesis Selection System for Medical Decision-Making," Proc. 8th Hawaii International Conf. on Systems.

Kulikowski, C.A. and N.S. Sridharan (1975) "Report on the First Annual AIM Workshop on Artificial Intelligence in Medicine." Sigart Newsletter No. 55, December 1975.

Kulikowski, C.A. (1976) "Computer-Based Consultation Systems as a Teaching Tool in Higher Education," 3rd Annual N.J. Conf. on the use of Computers in Higher Education, March 1976.

Kulikowski, C.A., S. Weiss, A. Safir, et al (1976) "Glaucomal Diagnosis and Therapy by Computer: A Collaborative Network Approach" Proc. of ARVO, April 1976.

Kulikowski, C.A., S. Weiss, M. Trigoboff and A. Safir (1976) "Clinical Consultation and the Representation of Disease Processes: Some AI Approaches," AISB Conferences, Edinburgh, July 1976.

Kulikowski, C.A., S. Weiss, A. Safir, (1976) "The CASNET Consultation System for Glaucoma: Development by Collaborative Computer Networking," Biosciences Communication Seminars, Aspen, Colorado.

Kulikowski, C.A., and A. Safir (1976) "Concepts of Computer-Based Modeling in Optics and Refraction," Proceedings Association for Computing Machinery Conference, Houston, Texas, pp. 110-114.

Kulikowski, C.A., S. Weiss and A. Safir (1977) "Computer Networking in the Development of a Glaucoma Consultation System" EDUCOM Fall Conference.

Kulikowski, C.A. (1977) "Representation of Knowledge for Reasoning in Medical Consultation," Panel Presentation, IJCAI, Boston.

Kulikowski, C. A (1977) "Problems in the Design of Knowledge Bases for Medical Consultation," The First Annual Symposium in Medical Care, Washington, D. C.

LeFaivre, R. and A. Walker (1975) "Rutgers Research Resource on Computers in Biomedicine, H," Sigart Newsletter No. 54, October 1975.

LeFaivre, R. (1976) "Procedural Representation in a Fuzzy Problem-Solving System," Proc. Natl. Computer Conf., New York, June 1976.

LeFaivre, R. (1977) "Fuzzy Representation and Approximate Reasoning," submitted to IJCAI-77, MIT.

Mathew, R., C. Kulikowski, and K. Kaplan (1977) "Multilevel Representation for Knowledge Acquisition in Medical Consultation Systems," Proc. MEDINFO Conference, Toronto.

Mauriello, D. (1974) "Simulation of Interaction Between Populations in Freshwater Phytoplankton," Ph.D. Thesis, Rutgers University 1974.

Safir, A., S. Podos, C. Kulikowski, S. Weiss, et al (1976) "Glaucoma Consultation by Computer," Scientific Exhibit E-22, Ophthalmology Transactions, American Academy of Ophthl. and Otolar., Vol. 81, No. 4.

Schmidt, C. (1972) "A comparison of source unidimensional, multidimensional and set theoretic models for the prediction of judgements of trail implication," Proc. Eastern Psych. Assoc. Meeting, Boston, April 1972.

Schmidt, C.F. and J. D'Addamio (1973) "A Model of the Common Sense Theory of Intension and Personal Causation," Proc. of the 3rd IJCAI, August 1973.

Schmidt, C.F. and A. Sedlak (1973) "An Understanding of Social Episodes," Proc. of Symposium on Social Cognition, American Psych. Assoc. Convention, Montreal, August 1973.

Schmidt, C.F. (1975) "Understanding Human Action," Proc. Theoretical Issues in Natural Language Processing: An Interdisciplinary Workshop in Computational Linguistics, Psychology, Artificial Intelligence, Cambridge, Mass., June 1975. Also appears as Computers in Biomedicine, TM-47, June 1975, Rutgers University.

Schmidt, C. (1975) "Understanding Human Action: Recognizing the Motives," Cognition and Social Behavior, J.S. Carroll and J. Payne (eds.), New York: Lawrence Earlbaum Associates, in press. Also appears as Computers in Biomedicine, TR-45, June 1975, Rutgers University.

Schmidt, C.F., N.S. Sridharan and J.L. Goodson (1976) Recognizing plans and summarizing actions. Proceedings of the Artificial Intelligence and Simulation of Behavior Conference, University of Edinburgh, Scotland, July 1976.

Schmidt, C. (1976) Understanding human action: Recognizing the plans and motives of other persons. In (eds. J. Carroll and J. Payne) Cognition and Social Behavior, Potomac, Maryland: Lawrence Earlbaum Associates, 1976.

Schmidt, C.F. and J.L. Goodson (1976) "The Subjective Organization of Summaries of Action Sequences," 17th Annual Meeting of the Psychonomic Society, St. Louis, 1976.

Schmidt, C.F., N.S. Sridharan, and J. L. Goodson (1977) "Plan Recognition: A Hypothesize and Revise Paradigm" Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass.

Schmidt, C. F, N. S. Sridharan, and J. L. Goodson (1978) "The Plan Recognition Problem: An Intersection of Psychology and Artificial Intelligence," Artificial Intelligence, Special Issue on applications to the sciences and medicine, to appear.

Sridharan, N.S. (1976) "The Frame and Focus Problems in AI: Decision in Relation to the BELIEVER System. Proceedings of the Conference on Artificial Intelligence and the Simulation of Human Behavior, Edinburgh, July 1976.

Sridharan, N.S. (1976) "An Artificial Intelligence System to Model and Guide Organic Chemical Synthesis, Planning in Chemical Synthesis by Computer, American Chemical Society Press, September 1976.

Sridharan, N.S. and C.F. Schmidt (1977) "Knowledge-Directed Inference in BELIEVER" in Pattern-Directed Inference, D. Waterman and F. Hayes-Roth (Editors) Academic Press (Scheduled for September).

Sridharan, N.S. and F. Hawrusik (1977) "Representation of Actions with Side-effects". In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Boston.

Srinivasan, C.V. (1973) "The Architecture of a Coherent Information System: A General Problem Solving System," Proc. of the 3rd IJCAI, August 1973.

Trigoboff, M. (1976) Propagation of Information in a Semantic Net," Proc. of the Conference on Artificial Intelligence and the Simulation of Behaviour, Edinburgh, Scotland, July 1976; updated version appears in CBM-TM-57, Dept. of Computer Science, Rutgers University, 1977.

Trigoboff, M. and C. Kulikowski (1977) "IRIS: A System for Propagation of Inference in a Semantic Network," Proc. Fifth International Joint Conference on Artificial Intelligence, Boston.

Tucker, S.S. (1974) Cobalt Kinetics in Aquatic Microcosms," Ph.D. thesis, Rutgers University.

Van der Mude, A. and A. Walker (1976) "Some Results on the Inference of Stochastic Grammars," abstract in Proc. Symposium on New Directions and Recent Results in Algorithms and Complexity. Dept. of Computer Science, Carnegie-Mellon University; Information and Control, (in press).

Van der Mude, A. and A. Walker (1978) "On the Inference of Stochastic Regular Grammars," Information and Control, (to appear).

Vichnevetsky, R. (1973) "Physical Criteria in the Evaluation of Computer Methods for Partial Differential Equations," Proc. 7th International AICA Congress, Prague, Sept. 1973; reprinted in Proc. of AICA, Vol. XVI, No. 1, Jan. 1974, European Academic Press, Brussels, Belgium.

Vichnevetsky, R., K.W. Tu, and J.A. Steen (1974), "Quantitative Error Analysis of Numerical Methods for Partial Differential Equations," Proc. 8th Annual Princeton Conference on Information Science and Systems, Princeton University, March 1974.

Walker, A. (1975) "Formal Grammars and the Regeneration Capability of Biological Systems," Journal Comp. and Syst. Sciences, Vol. 11, No. 2, 252-261.

Weiss, S. (1974) "A System for Model-Based Computer-Aided Diagnosis and Therapy," Parts I and II, Ph.D. Thesis, Rutgers University; also Computers in Biomedicine TR-27, Feb. 1974.

Weiss, S. (1976) "A System for Interactive Analysis of a Time-Sequenced Ophthalmological Data Base, Proceedings of Third Illinois Conference on Medical Information Systems.

Weiss, S., C. Kulikowski, and A. Safir (1978) "Glaucoma Consultation by Computer," Computers in Biology and Medicine, (in press).

Weiss, S., C. Kulikowski, A. Safir, "A Model-base for Computer-Aided Medical Decision-Making," Journal of Artificial Intelligence, (in press).

II. INTERACTIONS WITH SUMEX-AIM RESOURCE

During the past year we have continued to use the SUMEX-AIM resource for program development and testing, for communications among collaborators distributed in different parts of the country and for preparation and running of the AIM Workshop.

Among new projects developed in recent months are the TENEX SAIL system, which will also be used under the new 2050T system at Rutgers. The SUMEX-AIM system was also used to rewrite BRIGHT to work on TENEX, to debug TOPS-20 software, and to further develop the IRIS system. In another application, SUMEX-AIM was instrumental as a communications vehicle when the conversion of the Rutgers facility to the new 2050T system was being planned.

We continue to access SUMEX-AIM via TYMNET, and to a smaller extent via ARPANET. SUMEX-AIM played a key role in consolidating our network of collaborators in ophthalmology (ONET) and in providing the support needed for establishing a productive collaboration among the ONET investigators. It has also been most useful in communicating, planning and helping to set up the information pool for the third AIM Workshop.

Computing in the Rutgers Research Resource continues to be distributed between SUMEX-AIM and the RUTGERS-10. The two computers are providing complementary resources for our research and for our national collaborations. At present, the distribution of our computing is about 20 to 1 between RUTGERS-10 and SUMEX-AIM. Most of our work on SUMEX-AIM is done in INTERLISP (about 80% of our total connect hours) and the rest devoted mainly to communications and to limited program testing within ONET.

The SUMEX-AIM facility was used for demonstrations of AIM programs in first year classes and in second year seminars at the Rutgers Medical School: CASNET, MYCIN, INTERNIST and PARRY were accessed interactively in these classes and seminars. Another innovative use of SUMEX-AIM has been the collaborative development of the AI HANDBOOK, which is intended to provide a computer-based, network-accessible encyclopedia on Artificial Intelligence for the AIM community and guests. The AI HANDBOOK, was initiated by Dr. E. Feigenbaum and his students at Stanford. For the second year, a Rutgers graduate class, given by Dr. S. Amarel, worked on the Handbook and contributed several articles.

Finally, the SUMEX-AIM bulletin board continues to play an important role in communicating ideas and information on services among users. Since the MYCIN group at Stanford regularly posts summaries of meetings and other technical information on the MYCIN bulletin board, we have been able to keep track of their program and problems. This was particularly useful for our work on IRIS, where concepts close to the MYCIN CF formalism are being studied.

III. RESEARCH PLANS (8/78 - 7/81)

A. Long Range Project Goals and Plans

We are planning to continue along the main lines of research that we have established in the Resource to date, with emphasis on broadening our activities in the medical systems area. We are also planning to increase our participation in AIM dissemination and training activities, and to enhance the RUTGERS-10 computer in order to provide the support needed within the Resource and to increase the shared computing capabilities of the national AIM community.

B. Plans for Computing at Rutgers and at Stanford; development of the RUTGERS-10 facility as a node in the AIM network

By combining the NIH-recommended funding for computing over the renewal period of the Rutgers Resource (Dec. 1, 1977 to Nov. 30, 1980) and the Rutgers commitment for support of our research computing, the following plan for computing was developed and is now being implemented.

1. The current KI system at Rutgers was completely purchased in Dec. 1977. A new configuration was ordered to replace the present system in July 1978. The new system will be a KL-2050T with 512 K words of core and a TOPS-20 operating system.

2. The RUTGERS-10 computer is now controlled by the newly created Laboratory for Computer Science Research (LCSR) in which the Research Resource is administratively located; it is operated by a special CCIS group under a facilities management arrangement with LCSR. As Director of LCSR and PI of the Resource, Dr. S. Amarel has policy responsibility for the RUTGERS-10 facility vis-a-vis BRP and the University.

3. The user capacity of the RUTGERS-10 will be allocated as follows: (a) 55% to NIH grant activities - of this share, 3/4 will be allocated within the Resource for collaborative and core research, and 1/4 for the national

AIM community. [Thus, about 14% of user capacity of the RUTGERS-10 will be allocated to the AIM community.]; (b) 45% to Rutgers - this will be mainly devoted to computer science research (outside the NIH-supported Resource) and to advanced instruction in DCS. The user capacity is assumed to be 80% of total capacity, with the remaining 20% devoted to local systems activities and operations. Thus, an annual allocation of about 6,000 connect hours of KL (with about 1:60 compute to connect ratio) will be available to the AIM community starting in July 1978. About 20% of this will be needed for the AIM Workshop; the remaining suballocations will be devoted to national AIM users outside Rutgers. Based on projections of computing demand within the Rutgers Resource, the 1978 level of allocation may not be maintained in 2-3 years without enhancement of the KL configuration, although it is expected that at least 3,000 connect hours will be available in 1980.

4. The share devoted to the AIM community will be governed within the management framework of the existing national AIM committee structure - in the same manner that the national share of the SUMEX-AIM facility at Stanford is currently being governed. We propose to place special emphasis in the Rutgers AIM node on collaborative developments of knowledge-based systems in medicine. Also, since the TOPS-20 system (which represents the new line of DEC-supported operating systems) will be available on the Rutgers KL-2050T, it can provide the basis for technical experiments by SUMEX-AIM systems staff in preparation for future system changes in the AIM network. Such system support/planning activities for the AIM community should receive high priority in AIM usage on the RUTGERS-10.

This planning is based on the assumption that computing by investigators in the Rutgers Resource will continue to be distributed between the RUTGERS-10 and SUMEX-AIM, with the bulk of computing being done at Rutgers. We expect a demand level of about 2,000 connect hours per year on SUMEX-AIM - mainly for INTERLISP-based system developments, communications, special software developments and collaborative activities including AIM Workshops and the AI handbook.

E. FUNDING INFORMATION

1. Title of Grant or Contract:

   Rutgers Research Resource on Computers in Biomedicine

2. Principal Investigator(s): Saul Amarel
   Title and Institutional Affiliation: Professor and Chairman
   Department of Computer Science; and
   Director, Laboratory for Computer Science Research
   Rutgers University

3. Funding Agency: National Institutes of Health
   Biotechnology Resources Program
   Division of Research Resources

4. Grant or Contract Identification Number(s):

No. 2 P41 RR-00643

5. Total award term(s):

Three year renewal grant

Dates	Funding Amounts (DIRECT COSTS ONLY)
12/1/78-11/30/78	$505,823
12/1/78-11/30/79	522,005
12/1/78-11/30/80	437,064
	$1,464,892

6. Current Term(s), Dates, and Funding Amount(s):

12/1/77-11/30/78 $505,823

4.1.8 SIMULATION OF COMPREHENSION PROCESSES

Simulation of Comprehension Processes (SCP)

James G. Greeno and Alan M. Lesgold
Learning Research And Development Center
University of Pittsburgh

I. SUMMARY OF RESEARCH PROGRAM

The SCP project has only been on SUMEX for several weeks, so there is not too much to report yet. Even worse, these weeks have been mostly the period in which various professional meetings and other responsibilities have kept the investigators (Greeno and Lesgold) away from their work. Nonetheless, there is some progress to report and other information to be provided for this document.

Technical Goals

The goals of this project remain the simulation of young children's behavior in arithmetic and reading tasks, done in such a way that various levels of proficiency, from moderate dysfunction to expertise, can be modeled. In the reading work, the emphasis is on the interaction of several components of the reading process: word recognition, sentence parsing, and relating new sentence content to what has already been understood. The project begins from the empirical basis of a large number of studies showing that various word processing skills are less developed in deficient readers. Lesgold and Perfetti (in press) have discussed this evidence and have suggested that poorly developed word processing skills use too much of a limited processing capacity, thereby indirectly hampering comprehension skills that may otherwise be operating at acceptable levels.

Regrettably, the evidence to support this point of view has been largely correlational. To support a more detailed and empirically verified specification of the sources of reading dysfunction, we have adopted the strategy of combining several converging empirical investigations with a simulation of the reading process that can more clearly specify the convergence of evidence. Thus, we are conducting a number of empirical studies of comprehension processes, have reviewed a large body of evidence on word recognition (e.g., Perfetti & Lesgold, in press; Lesgold & Perfetti, in press, 1977), and are conducting a longitudinal study of changes in children's word processing and reading abilities throughout the course of primary-grades reading instruction.

In our work on arithmetic, we are continuing the development of a model of comprehension processes related to elementary arithmetic concepts. The model simulates a process of solving simple word problems, constructing an integrated representation of a problem using schematic knowledge about quantitative relationships (Greeno, 1978). We are conducting a developmental study to determine the degree to which semantic and linguistic factors, rather than arithmetic knowledge, are responsible for children's difficulty in solving problems at early grade levels.

Medical Relevance and Collaboration

The range of ability levels we will be dealing with in arithmetic and reading includes children who are below average and, in the case of reading, some children who are classed as "learning-disabled." By providing a framework within which the effects of differing levels of skill acquisition can be understood, we hope to eliminate the spurious use of vague medical categories such as "minimal brain damage," etc., and thus more clearly delimit the cases in which there is a real medical problem in a child whose achievement in math and reading is poor. Toward this end, we have collaborated with Isabel Beck of the Learning Research and Development Center in doing comparisons of our data from "normal" children in reading with her data on "LD" children. The comparisons should, at a later stage of the simulation work, play a role in the design of an overall model for reading.

Progress Summary

The few weeks that we have been on SUMEX have been spent in learning the system and in interacting with the ACT project (see below) to develop a complete understanding of the resources that can be borrowed from them rather than reinvented. The empirical progress on the reading work is described in the publications listed below and in a paper presented on March 30, 1978, at the annual meeting of the American Educational Research Association. Recent progress on the arithmetic work will be reported in papers to be presented at the May meetings of the Midwestern Psychological Association, as well as in Greeno's paper cited below.

List of Relevant Publications

Greeno, J.G. Some examples of cognitive task analysis with instructional implications. Paper presented at the ONR/NPRDC Conference on Aptitudes, Cognition, and Instruction, San Diego, March, 1978.

Greeno, J.G. A study of problem solving. In R. Glaser (Ed.), Advances in instructional psychology. Hillsdale, NJ: Erlbaum, in press.

Heller, J.I., & Greeno, J.G. Semantic processing in arithmetic word problem solving. Paper presented at Midwestern Psychological Association, May, 1978.

Lesgold, A.M., and Perfetti, C.A. Interactive processes in reading. Discourse Processes, in press.

Perfetti, C.A., & Lesgold, A.M. Discourse processing and sources of individual differences. In P. Carpenter & M. A. Just (Eds.), Cognitive processes in comprehension, Hillsdale, NJ: Erlbaum, 1977.

Perfetti, C.A., & Lesgold, A.M. Coding and comprehension in skilled reading. In L.B. Resnick & P. Weaver (Eds.), Theory and practice of early reading, Hillsdale, NJ: Erlbaum, in press.

Riley, M.S., & Greeno, J.G. Importance of semantic structure in difficulty of arithmetic word problems. Paper presented at Midwestern Psychological Association, May, 1978.

Funding support status

A. ONR
1. Analysis of Formal and Informal Reasoning in Problem Solving
2. James Greeno, Research Associate and Professor of Psychology, Learning Research and Development Center, University of Pittsburgh..
3. Office of Naval Research
4. Contract Number N0014-78-C-0022
5. October 1, 1977 through September 30, 1980; $198,000.
6. October 1, 1977 through September 30, 1978; $62,616.

B. NIE
1. Adaptive Education
2. Robert Glaser (James Greeno and Alan Lesgold are both Research Associates of the Learning Research and Development Center and are funded out of this grant).
3. National Institute of Education
4. Grant Number OB-NIE-78-0115.
5. December 1, 1977 through November 30, 1978; $3,800,000.
6. December 1, 1977 through May 30, 1978; $700,000.

(N.B. the current level of funds expended on the specific research projects of Greeno and Lesgold from this funding source are $70,140 and $55,760 per annum, direct costs, respectively. In addition, Greeno and Lesgold use substantial computer resources and other technical resources for experimentation and data analysis that are funded from the overall grant but not charged to their specific budgets.)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

Collaborations and medical use of programs via SUMEX

No such collaborations have yet been implemented.

Sharing and interactions with other SUMEX-AIM projects

We have been in continual touch with the members of the ACT project, primarily because our simulations will be developed within the ACT system. However, these interactions have been bilateral in effect. As we have asked questions, details of ACT have become better understood and better implemented.

Critique of resource management

We have enjoyed clearly adequate resources so far. One suggestion to be made is that national users be regularly informed of the telecommunications costs associated with access to SUMEX-AIM so that responsible decisions can be made regarding use of the system.

III. RESEARCH PLANS

Long range project goal and plans

As work on modeling arithmetic knowledge proceeds, we hope to be able to develop theoretical analyses of the process of acquisition of elementary mathematics. In this work, we anticipate further collaboration with the ACT project, in which recent efforts have been focused on mechanisms of learning. Preliminary discussions of collaborative work on the acquisition of strategic knowledge in geometry problem solving, such as that described in Greeno (in press).

The goals for reading are similar, but tend to more heavily emphasize development of models of components of the reading process that are able to easily characterize the wide range of individual differences in reading that are found in the U.S. The work will involve continual alternation between empirical investigations and use of SUMEX-AIM for modeling. We also anticipate substantial interactions with Anderson and Kline of the ACT project to continue over the life of the project.

Justification and requirements for continued SUMEX use

The primary justification for our access to SUMEX-AIM is that it is the "home" of both the versions of ACT that we need for our work and the ACT researchers with whom we can benefit from continual interaction, exchange of developing simulation programs, etc. An additional important justification is that we do not presently have access to a computer system that can support ACT's more useful versions, nor is such a facility likely to become available. A final benefit is our ability to interact with other researchers through ARPAnet. We anticipate that other workers will be able to run demonstrations of our simulations, as they develop, through the SUMEX-AIM guest arrangements.

At the present time, our allocations of disk space within SUMEX are adequate, though it is conceivable that before a year passes we will need a small increment in space. If so, a request will be made to the Executive Committee at that time.

Needs and plans for other computer resources beyond SUMEX-AIM

We make substantial use of the computer systems at the University of Pittsburgh for all the data analysis and on-line experimental control services that the empirical side of our research requires. In addition, a more primitive version of ACT has been installed on the local system here. Thus, our use of

SUMEX-AIM is limited to the specific simulation needs which cannot otherwise be accomplished.

Recommendations for future community and resource development

We do not have enough experience on SUMEX-AIM to make recommendations yet.

4.2 STANFORD PROJECTS

The following group of projects is formally approved for access to the Stanford aliquot of the SUMEX-AIM resource. Their access is based on review by the Stanford Advisory Group and approval by Professor Lederberg as Principal Investigator. As noted previously, the DENDRAL project was the historical core application of SUMEX. Although this is described as a "Stanford project," a significant part of the development effort and of the computer usage is dedicated to national collaborator-users of the DENDRAL programs.

4.2.1 AIHANDBOOK PROJECT

Artificial Intelligence Handbook

E. A. Feigenbaum
Stanford Computer Science Department

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The AI Handbook is a compendium of knowledge about the field of Artificial Intelligence. It is being compiled by students and investigators at several research facilities across the nation. The scope of the work is broad: Two hundred articles cover all of the important ideas, techniques, and systems developed during 20 years of research in AI. Each article, roughly four pages long, is a description written to be suitable both for the non-AI specialist and students of AI. Additional articles serve as Overviews, which discuss the various attempts within a subfield, the issues, and the problems.

There is no comparable resource for AI researchers and other scientists who need access to descriptions of AI techniques like problem solving or parsing. The research literature in AI is not generally accessible to outsiders. And the few textbooks that exist are not nearly broad enough in scope to be useful to a scientist working primarily in another discipline who wants to do something requiring knowledge of AI. Furthermore, we feel that some of the Overview articles are the best critical discussions available anywhere of activity in the field.

To indicate the scope of the Handbook, we have included an outline of the articles as an appendix to this report (see Appendix I on page 217).

B. Medical Relevance and Collaboration

The AI Handbook Project was undertaken as a core activity by SUMEX in the spirit of community building that is the fundamental concern of the facility. We feel that the organization and propagation of this kind of information to the AIM community, as well as to other fields where AI is being applied, is a valuable service that we are uniquely qualified to support.

C. Progress Summary

Because our objective is to develop a comprehensive and up-to-date survey of the field, our article-writing procedure is suitably involved. First drafts of Articles are reviewed by the staff and returned to the author (either an AI scientist or a student in the area). His final draft is then incorporated into a Chapter, which when completed is sent out for review to one or two experts in that particular area, to check for mistakes and omissions. After corrections and comments from our reviewers are incorporated by the staff, the manuscript is edited, and a final computer-prepared, photo-ready copy of the Chapter is generated.

We expect the Handbook to reach a size of approximately 700 pages. By that estimate, we are now 1/4 finished and expect to reach the halfway point during the Summer of 1978, since a good deal of the material outstanding has already been submitted in first draft form. Volume I of the Handbook, which will include the material completed in Summer 1978, will cover AI research in Heuristic Search, Representation of Knowledge, AI Languages, Natural Language, Speech Understanding, Applications-oriented AI research, and Automatic Programming. It may also include the chapters on Information Processing Psychology, Learning and Planning. Researchers at Stanford University, Rutgers University, SRI International, XEROX PARC, RAND Corporation, and others have already contributed articles or reviewed chapters.

D. List of Relevant Publications

Volume I of the AI Handbook will appear in preliminary from as a Stanford Computer Science Technical Report in the Fall of 1978.

E. Funding Support Status

The Handbook Project is currently supported under the Heuristic Programming Project contract with the Advance Research Projects Agency of the DOD, contract number MDA 903-77-C-0322, E. A. Feigenbaum, Principle Investigator. During the next year, SUMEX core research funds will provide partial personnel support for approximately 1.4 FTE.

II. <u>INTERACTIONS WITH SUMEX-AIM RESOURCE</u>

A. Collaborations and medical use of programs via SUMEX

We have had a modest level of collaboration with a group of students and staff at the Rutgers resource, as well as occasional collaboration with individuals at other ARPA net sites.

B. Sharing and interactions with other SUMEX-AIM projects.

As described above, we have had moderate levels of interaction with other members of the SUMEX-AIM community, in the form of writing and reviewing Handbook material. During the development of this material, limited arrangements have been made for sharing the emerging text. As final manuscripts are produced, they will be made available to the SUMEX-AIM community both as on-line files and in the hardcopy, published edition.

C. Critique of Resource Management

Our requests of the SUMEX management and systems staff, requests for additional file space, directories, systems support, or program changes, have been answered promptly, courteously and competently, on every occasion.

III. RESEARCH PLANS (8/78 - 7/81)

A. Long Range Project Goals

The following is our tentative schedule for completion and publication of the AI Handbook:

May through August, 1978 - Various Chapters of Volume I will be available as Stanford Computer Science Technical Reports, distributed to the AI community for peer review.

Fall 1978 - After revision reflecting criticism of the reviewers, material for Volume I will be submitted for publication.

Fall 1978 through Spring 1979 - Development and refinement of material in remaining Chapters, Volume II.

Summer 1979 - Completion and publication of Volume II.

B. Justifications and requirements for continued SUMEX use

The AI Handbook Project is a good example of community collaboration using the SUMEX-AIM communication facilities to prepare, review, and disseminate this reference work on AI techniques. The Handbook articles currently exist as computer files at the SUMEX facility. All of our authors and reviewers have access to these files via the network facilities and use the document-editing and formatting programs available at SUMEX. This relatively small investment of resources will result in what we feel will be a seminal publication in the field of AI, of particular value to researchers, like those in the AIM community, who want quick access to AI ideas and techniques for application in other areas.

C. Your needs and plans for other computational resources

We use document preparation facilities (the XEROX Graphics Printer) at the Stanford AI Laboratory.

D. Recommendations for future community and resource development

None.

4.2.2 DENDRAL PROJECT

RESOURCE RELATED RESEARCH -- COMPUTERS AND CHEMISTRY: THE DENDRAL PROJECT

Carl Djerassi, Principal Investigator
Professor of Chemistry
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

OVERVIEW OF RESEARCH ACTIVITIES

In this first year of a three year renewal, substantial progress was made on every major item in the renewal proposal. The most obvious facets of this interdisciplinary work on computers and chemistry are research, engineering and applications. On the research side, the computer programs have grown in both chemical and computer science sophistication. On the engineering side, the programs have been made faster and easier to use. On the applications side, the programs have been used by chemists working on biomedical problems at Stanford and elsewhere as aids in their own research (see ref 4).

STRUCTURE ELUCIDATION PROGRAMS

Stereochemistry in CONGEN

The set of computer programs developed at Stanford as tools for molecular structure elucidation are being enhanced by the addition of 3-dimensional structural information. The programs can now deal with some basic geometrical properties of molecules that are essential for understanding their biological significance. Research progress this year has resulted in extensions that allow computation of stereoisomers (alternative structures differing in 3 dimensions but having identical connections among atoms). Thus geometrical variations on structural hypotheses can be presented as well as topological variations.

The proposed first stage in this effort was to write a program which was capable of recognizing the configurational stereochemical features of a molecule and generate all the possible stereoisomers based on these features. This program has been written and interfaced to an experimental version of CONGEN, and is described in detail below. The proposed second stage in this effort is to modify this program to permit generation of stereoisomers which satisfy certain constraints, much as the existing CONGEN program constrains the generation of topological isomers. This ongoing effort is discussed in the section on future plans.

Example

The structure is 3-6-dimethyl-4-octene, a simple hydrocarbon which exhibits double bond and configuration stereochemistry. It has only a small number of stereoisomers because the symmetry of the molecule reduces the total number that are possible.

```
3-6-dimethyl-4-octene

              4
              |
        1-2-3-5=6-7-9-10
                  |
                  8


THERE ARE 6 STEREOISOMERS

        0 1 1 -1
        1 0 1 1
        2 0 1 1
        4 1 1 -1
        5 0 1 1
        6 0 1 1
```

The first number on each row is the canonical label for each stereoisomer. The correspondence is:

```
0  R-S-trans
1  S-S-trans
2  R-R-trans
4  R-S-cis
5  S-S-cis
6  R-R-cis
```

The second number on each row tells whether this particular stereoisomer is achiral (1) or has an enantiomer (0). Enantiomeric pairs are listed on consecutive rows. The final two numbers on each row indicate the symmetry group of each stereoisomer. Those with 1 1 have rotational symmetry and those with 1 -1 have a plane of symmetry.

Future Plans

The following features (at least) will be added to the existing program:

1) Designations of stereocenters as either R or S based on constitutional priorities only. This will be for aid in interpretation only as these designations are not useful internally to the program.

2) Recognition of cis and trans double bonds for the same reason.

3) Stereoisomer output which is interpretable and compatible with character terminal output. This will most likely be done in conjunction with the existing drawing program. The compatibility with character based terminals is a strength of CONGEN at present.

4) Versatility in the handling of the stereochemistry of atoms other than carbon. In particular there should be a choice as to whether a nitrogen atom is thought to be able to invert freely.

The second stage of the development in this effort is to give CONGEN the ability to constrain stereoisomer generation. The algorithm of the generator was designed so that a number of useful constraints, particularly concerning relative stereochemistry between stereocenters can be applied prospectively. That is, the undesired stereoisomers would not be generated. Other constraints, such as those which involve the symmetry of the stereoisomers can be applied during the generation. Finally, there will certainly be some constraints which have to be applied after generation.

Constraints Interpretation

The area of automatic interpretation of constraints in CONGEN structure elucidation problems is interesting and important for two reasons: 1) we want to free the chemist as much as possible from having to understand CONGEN's method of building structures; and 2) problems can be solved much more efficiently if CONGEN can perform some preliminary examination of them and find an alternative, efficient way to solve the problem. Our first efforts in this direction have resulted in what we call the "GOODLIST interpreter", which is designed to make more efficient use of information about <u>required</u> (GOODLIST items plus Superatoms) structural features of an unknown molecule.

Introduction to Method

It is characteristic of structure elucidation based on data from physical and chemical methods that much structural information is redundant. Physical methods, for example, are frequently complementary. One technique provides structural information which can be used to elaborate information gathered by another. The collection of partial structures present in an unknown derived by such methods frequently contain atoms or groups of atoms shared among two or more partial structures. Chemists must take this into account when considering how the partial structures might fit together to yield the structure of an unknown compound. We are developing a method which is designed to demonstrate the task of determining potential overlaps in partial structures.

Stated in the simplest terms, the method should translate the constraints on desired structural features, or GOODLIST constraints, into new sets of partial structures which incorporate the features at the <u>beginning</u> of the structure generation procedure.

Future Directions

Initially, the GOODLIST substructures specified as constraints will be incorporated automatically at the beginning of the problem as described above. Within a short time, the method of specification of a problem will be changed to include only the empirical formula together with inferred partial structures without regard to overlaps, leaving to the program the task of determining those overlaps and specifying the set of problems to solve.

Incorporation of BADLIST (undesired structural features) substructures in the procedure is a necessary next step. Subsequently we will attack the problem of discerning constraints which are implied by the input data, including detection of unclear or ambiguous statements about a structure. The constraints interpreter should be capable of a dialog with the chemist using CONGEN to clarify such points prior to structure generation.

Experiment Planning Program

Now that CONGEN gives us the capability of constructing all plausible candidates under an initial set of constraints, the next problem is to provide the chemist with some assistance in rejecting incorrect candidates and focussing on the correct structure. This process must involve the examination of the candidates to determine their common and unique features, and the designing of experiments to differentiate among them.

The initial work on this problem has begun by providing a new function, the EXAMINE function, which gives a chemist the ability to survey sets of structures for particular combinations of substructures, ring-systems etc. This function has now been incorporated into the CONGEN program. EXAMINE allows structures to be segregated on the basis of combinations of (desired or undesired) structural features. For example, EXAMINE can be used to segregate structures which possess feature A or B, or generally, any arbitrary Boolean expression of relationships among structural features.

More elaborate functions for automatically identifying discriminating features in sets of structures are being developed. Currently, these experimental routines (contained within the "PLAN" program) can be used to analyze functionality, or to identify differences in the ways that superatoms have been imbedded in structures. These routines will shortly be capable of exploiting a simplified library of chemical/spectral tests for particular substructural features; this will allow the program to identify possible discriminating experiments.

The Reaction Chemistry Program

During the past year we have made progress in developing the reaction chemistry program, REACT, into a working tool for laboratory chemists.

REACT is designed to carry out representations of chemical reactions on representations of chemical structures. Reactions, defined by the chemist using the program, are carried out in the synthetic direction as opposed to the retrosynthetic direction of programs for computer-aided synthesis. (6) In structure elucidation problems, the set of structures undergoing reaction is the current set of candidate structures for an unknown. It is clear, however, that the program can also be used effectively in following reactions of a single, known compound participating in a complex sequence of reactions. For example, we showed (ref 22) that CONGEN together with REACT provides a convenient method for studying acid catalyzed rearrangements such as the conversion of tetrahydrodicyclopentadiene to adamantane. In that example, the complete set of isomers was generated by CONGEN. Subsequently, a one-step reaction carried out on each isomer afforded the complete rearrangement graph.

During the past year the structure of the program was revised significantly to include commands and internal operations which more closely parallel laboratory procedures. The new version has been described briefly and some applications of REACT to mechanistic problems have been discussed (ref 24).

Whether applied to mechanistic studies or to structure elucidation problems, REACT's soperatins are based on procedures such as (1) carrying out reactions in a variety of ways, (2) separation of products, (3) labelling of "flasks" for products, (4) carrying out further reactions on products, and (5) testing the contents of flasks against structural constraints.

Structural information obtained in the laboratory on the content of a product flask represents constraints not only on the identity of the product, but also on the identity of the precursor and its precursor and so forth throughout an entire reaction sequence. REACT allows structural statements to be made as constraints on the contents of any flask in a reaction tree. The program then determines automatically the structural implications of the constraint throughout the reaction sequence.

Mass Spectral Prediction and Ranking

Predicting Spectra Using MSRANK and the Half-Order Theory

The MSRANK program has been incorporated as part of CONGEN, but is not yet available for general use. During the past year we have been giving the program some extensive tests to determine its scope and limitations. We have studied the following classes of compounds (all closely related to current research problems): 1) marine sterols; 2) substituted pregnanes; 3) aliphatic and aromatic esters; and 4) macrolide antibiotics.

MSRANK is a powerful filter for eliminating from further consideration structures which cannot yield the observed mass spectrum for an unknown by "reasonable" fragmentation pathways. The greater the structural diversity of isomeric candidates for an unknown, the better the performance of MSRANK in focussing in on the correct structure. When the structures are quite similar, for example when they have been constructed from the same set of superatoms and few remaining atoms, the ranking by MSRANK is quite similar (as one might expect). When this situation occurs, the chemist must still consider the top 10 - 50 percent of the structures as possibilities, depending on the distribution of scores.

---

(6) E.J. Corey and W.T. Wipke, Science 166, 178 (1969).

We have added an explanation feature to MSRANK. Upon request the program prints a list of peaks in the observed spectrum which have different "reasonable" explanations. for different candidate structures. Based on this information the chemist can accept the ranking or change the parameters which define his theory of fragmentation to obtain a different ranking. This procedure helps detect and reduce the plausibility of "nonsense" fragmentation processes.

Prediction Using Fragmentation Rules Supplied by Chemists

When the candidate structure is known to belong to a previously investigated class of compounds, then we can use additional information to predict a more precise mass spectrum. This information is in the form of specific fragmentation rules. These rules are described by a subgraph, a break (or cleavage) and related hydrogen or neutral transfers, intensity ranges associated with rules and a parameter describing the confidence in a rule. We are working on a program which allows the user to enter rules defining his theory of mass spectral fragmentation.

The next step is to explore ways to compare a predicted and an observed spectrum. We are experimenting with different ranking functions (see section 4) and developing a program which will allow the user to define in a simple mathematical equation his individual ranking function. The problem of ranking candidate structures based on spectrum comparison is closely related to the problem of library search. In our case, however, we do not have authentic spectra of our structural candidates in most instances. The density of a predicted spectrum for a candidate is quite low because we do not attempt to predict the complete spectrum. Rather, we predict major fragmentations. This fact must be taken into account in designing a function to rank candidates based on comparison of their predicted spectra to the spectrum of the unknown.

Molecular Ion Determination

The original MOLION program (7) was based upon the postulate: "There exists at least one SECONDARY LOSS in a spectrum that will match a PRIMARY LOSS from the molecular ion irrespective of whether the molecular ion is present in the spectrum."

Given this postulate, then one method of generating candidate masses for a molecular ion (M+) is to identify all possible secondary losses apparent in a spectrum, and then to add each of these losses to the masses of those ions observed in the high mass region of the spectrum. This, together with some refinements, was the basis of the original MOLION program.

There are, however, a number of problems with the algorithm used in MOLION. The most crucial problem is that the algorithm requires good spectra! Impurities such as column bleed or co-eluting minor components can result in ions that would

---

(7) R.G.Dromey, B.G.Buchanan, D.H.Smith, J.Lederberg and C.Djerassi. "Applications of Artificial Intelligence to Chemical Inference. XIV. A General Method for Predicting Molecular Ions in Mass Spectra." Journal of Organic Chemistry, 40, 770 (1975).

constitute bad losses --- causing the rejection of distinct and well supported molecular ions recorded in the spectrum. Further, the program did not allow the user to modify the "bad loss" set, nor to have access to the molecular ion scoring mechanisms. These scoring mechanisms incorporated a considerable measure of class dependency. Thus when testing a candidate M+, the program could modify the score associated with the M+ by the intensity combination formula: e.g. a mass difference of 101amu between the candidate M+ and an observed ion resulted in a 1.8 times increase in that M+'s score whereas a mass difference of 2 or 16 reduced the score by 85 per cent and a difference of 44, 56, 60 or 72 reduced the score by 25 per cent.

In devising the new version of the molecular ion program, an attempt has been made to recognize and overcome some of these problems. We have made modifications suggested by the above considerations. The resulting program is quite successful and has been incorporated into our sequence of programs for analysis of combined gas chromatographic/mass spectrometric (GC/MS) data.

CONGEN Improvements

During the past year many improvements have been made in the version of CONGEN available for outside use. These improvements allow the user more flexibility and range in the use of existing commands. Further, some new commands have been created which increase the power and utility of CONGEN. The program has become easier to use and more robust. Finally in almost every subsection of the program the user can inspect the computation as it proceeds. This means that fewer long, wasteful computations will be performed. Some of the major improvements are described briefly below.

Error Detection in Substructure Definition

We provide extensive error checking on structural fragments (substructures) defined by the user of CONGEN. If the chemist does not choose to fix the errors, we warn very clearly that the results will be unpredictable or erroneous. We allow the chemist to go ahead on the philosophy that he may have a perfectly good reason for doing what seems to us to be nonsense. We have concentrated our efforts on mistakes which we have observed when other chemists use CONGEN. Moreover, this error checking will serve to reduce substantially the errors made by chemists using the program.

Depth-First Imbedder

The IMBEDDER program was completely rewritten. Four major improvements were implemented and the efficiency of almost all of the different subsections was improved. First, the method of computation was changed from breadth first (all structures delivered at the same time) to depth first (the structures delivered one at a time as they are created). The chemist can now check the computation as it proceeds by using the cntrl-S and cntrl-I features. Use of either feature to interrupt the computation and examine results often will allow the chemist to see that a certain computation is much larger than he anticipated and to stop it before computer time is wasted.

Second, all the constraints testing during imbedding is now done in the SAIL portion of CONGEN and structures violating the constraints are not returned. Previously all structures were returned to the LISP portion of CONGEN before any constraints checking and subsequent pruning were done. This new approach represents a real gain in efficiency because these programs run much faster in SAIL than they do in LISP.

Third, the canonicalization routines were rewritten. New algorithms were found and the process of assigning a canonical number to a structure is now much less costly in terms of time. Further, two structures which are aromatically equivalent are now given the same key (related to its canonical number). Since many different parts of CONGEN use the canonicalizer this resulted in a gain in efficiency for all of them.

Fourth, a change was made so that any number of superatoms can be imbedded at one time. This means when large numbers of superatoms need to be imbedded the chemist can in one set of commands perform the entire task, rather than the more time consuming approach of one-at-a-time. However, the chemist can still choose for reasons of efficiency to imbed a single superatom when special tests on the environment of that superatom are required. This also provides the opportunity for large, multiple imbeddings to be done in batch mode. (The batch command was rewritten so that it would accept multiple superatoms.) The large batch job is then run after midnight when the load average is low.

EDITSTRUC Changes

The RENUMBER command was added to give the chemist flexibility in choosing schemes of numbering the atoms in the structure. There have been internal changes made to the editstruc commands BRANCH, LINK, CHAIN, and DELATS. All involve the method of numbering atoms. It is now possible to create a substructure which has "gaps"(missing numbers) in its numbering to atoms. These changes necessitated some further changes in the routines which prepare and send structures to the IMBEDDER in CONGEN.

BATCH

The BATCH command was rewritten to take advantage of the fact that the new imbedder can accept any number of superatoms to be imbedded. As the system load continues to increase BATCH will become a more attractive option.

CONGEN Reprogramming

We have been investigating the reprogramming of CONGEN into an Algol-like language. The goals of reprogramming are threefold: first, to unify the program into a single language which can be used on a variety of computer systems; second, to begin to compact the program into a manageable, cost-effective size for current time-sharing systems; and third, to improve typical runtimes for CONGEN so that it becomes a more attractive means for scientists to solve structure elucidation problems. A version of CONGEN which fulfills these goals would be useful on a variety of computer systems and could be exported to many different chemical and biochemical laboratories.

THEORY FORMATION PROGRAMS - Meta-DENDRAL

Incremental Learning

In order to allow applying the Meta-DENDRAL program (ref 3) to a wider range of chemically interesting problems, we have begun to remove one of the most important current program limitations - its inability to add piecewise to what it has learned. Meta-DENDRAL must currently process all training data at once, producing a set of rules which cover that data.

Since the amount of training data processed strongly influences the reliability of the learned rules, training on arbitrarily large data sets will allow Meta-DENDRAL to form more accurate rule sets than currently feasible.

Modifying Existing Rules

Rules must be modified so that they become consistent with the new data while remaining consistent with previous data as well. In short, the method involves storing along with each rule a summary of alternate acceptable versions of the rule (those with the same evidential support in the observed training data). The summary of all acceptable versions of a given rule, referred to as the version space (ref 14) of the rule, is useful for a number of tasks associated with rule learning, including incremental learning.

Version spaces provide an explicit representation of the space of all alternate versions of a given rule - i.e. those which cannot be disambiguated by the currently observed training data. As such, version spaces will allow Meta-DENDRAL to reason more thoroughly with the choice among alternate rule versions.

Current Status and Future Work

The incremental learning ability for Meta-DENDRAL is almost fully implemented, but as yet remains untested. Routines for defining and modifying rule version spaces are implemented, as well as the ability to filter out training data explained by a rule set. The major unimplemented portion of the incremental learning scheme is the process for merging new rules into the evolving rule set. The chief issue here is deciding when and how to chose among or merge new rules which are similar to existing rules. We expect to complete implementation and initial testing of the incremental learning ability during 1978.

Among issues associated with the version space approach which we expect to explore during the current grant period are the following:

1) Intelligent selection of new training data from examination of partial results.

2) Applying chemical plausibility information to select a "best" rule version from among those contained in the version space.

3) The extension of current methods for dealing more completely with noisy and ambiguous training data.

4) The use of version spaces for merging similar rules.

New Capability To Emphasize Discriminatory Power

One important intended use of rules formed by Meta-DENDRAL is the prediction of mass spectra for use in structure elucidation: Predicted spectra for a set of candidate structures are compared by computer with the mass spectrum observed for an unknown compound, and on this basis the candidates are ranked according to plausibility. The ability of rules, in this context, to differentiate correctly among candidate hypotheses is called their "discriminatory power." Since the selection criteria previously used by Meta-DENDRAL during the various stages of rule formation did not necessarily correlate with high discriminatory power, it was decided to provide the program with the option of directly emphasizing discriminatory power during rule formation, in order to maximize the usefulness of the resulting rules for purposes of structure elucidation.

This addition to Meta-DENDRAL has now been designed and implemented. The general method employed by the the new option is as follows. Observed mass spectra of the training molecules are analyzed prior to rule generation to determine how diagnostic the various observed peaks are, within the training set, of the molecules that show them. This information is then used during rule formation to compute a measure of discriminatory power for emerging rules. This measure is used, in combination with other criteria, to guide the search during rule generation, and to control the modification and selection of rules during the later phases of processing.

Preliminary testing of this new rule formation scheme on the monoketoandrostanes produced rules of considerably greater discriminatory power within that family than had been produced in earlier work with Meta-DENDRAL, even though the training set used was only half as large as that used earlier. This "discrimination option", now integrated with the new template-processing capability, is currently being further tested on a group of aromatic esters to determine whether the rules formed are consistent with what is known about the fragmentation modes of those molecules, and whether the rules have significant discriminatory power outside the training set used to form them.

Data Selection Program

Good inductive generalizations depend on variety in the data set. This is no less true in the context of rule formation by Meta-DENDRAL. Whether the goal is to discover rules of high generality or high discriminatory power, one's chances of achieving this goal appear to increase with increasing variety of training instances. This suggests that it would be useful to have a data selection program that would select the subset of the potential training molecules which has the greatest variety, in some appropriate and well-defined sense. A preliminary version of such a program has been implemented, and experiments with it will soon be underway.

Feedback Loops

The RULEGEN program is capable of accepting previously defined rules as a means of filtering the evidence obtained from INTSUM before the evidence is used for rule formation. As well as providing a convenient and natural feedback mechanism for the program, this facility also allows rules obtained from other sources to be used to reduce the space which the program must examine to find rules for a given set of data. In this manner, the program is able to focus attention on evidence which is not already explained by any of the rules which it is given. Tests are in progress to determine the limitations of this approach.

Stability Rules in INTSUM and RULEGEN

The programs have been generalized to allow the analysis of the mass spectral data from the point of view of determining rules about stable bonds, i.e., lack of fragmentation in a molecule as well as fragmentation. Just as peaks are evidence of fragmentation in a structure, absence of peaks is evidence that certain fragmentations have not occurred.

The programs are now capable of examining the original data from either point of view and proposing rules of behavior of the molecules from that point of view. Further work remains to be done to carry this generality through the processing performed in RULEMOD and then in conducting experiments to determine the usefulness of stability analysis.

Carbon-13 Work

The work described in this section was accomplished in conjunction with work on structure elucidation and theory formation programs (sections 2 and 4).

Carbon-13 nuclear magnetic resonance (CMR) has developed into an important tool for the structural chemist. A natural abundance CMR spectrum which is fully proton decoupled consists of a number of sharp peaks which correspond to the resonance frequencies in an applied magnetic field of the various types of carbon atoms present. A C-13 shift is the amount an observed peak is shifted from that of a reference peak, usually tetramethylsilane (TMS).

During the past year we continued work on an extension of Meta-DENDRAL which allows the program to form rules in the domain of CMR spectroscopy. We also wrote a second program which applies CMR rules to structure elucidation problems. Rules generated from a combined set of paraffins and acyclic amines have been used to successfully identify the C-13 NMR spectra of molecules not in the training set data. The introduction of a limited set of stereochemical terms to the rule generation procedure demonstrated the feasibility of extending the method to more complicated systems. A description of the rule formation and structure elucidation programs is given in (ref 16). Results are presented there for the combined set of paraffin and acyclic amines, as well as for a combined set of trans decalins and monohydroxylated androstanes.

Molecular structure elucidation is accomplished by our program by selecting a shift (peak) in the observed spectrum, then finding the rules which are possible explanations for this shift. The rules selected postulate partial substructures which might be in the molecule. These substructures are then assembled jigsaw puzzle fashion to construct the final molecule. Constraints stemming from both the observed spectrum and information associated with each rule are used to constrain the process so that only "reasonable" structures will be considered.

The structure elucidation program has been run on several test cases using unknown paraffin and acyclic amine spectra with reasonable success. This program is described in detail in (ref 16).

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

Use of CONGEN via SUMEX

The following individuals were among those listed in the DENDRAL annual report as persons who use CONGEN or who have requested information about access.

Dr. Peter Gund of Merck, Sharpe and Dohme Laboratories

Professor Richard E. Moore of the University of Hawaii

Dr. Jean-Claude Braekman of the University of Brussels

Dr. Martin Huber, a postdoctoral fellow in Professor Wipke's SECS group

Professor Kurt Mislow of Princeton University

Professor Weiss, head of the Department of Chemistry at Northeastern University.

Dr. Stan Lang of Lederle Labs' Infectious Disease Research Section

Dr. Leon Goldman and Dr. Babu Venkataraghavan, also at Lederle

Professor E.J. Eisenbraun of Oklahoma State University

Dr. David Pensak of Dupont in Wilmington, Delaware

Dr. Milton Levenberg of Abbott Laboratories

Kent Morrill of Tennessee Eastman

Dr. Gretchen Schwenzer of Monsanto

Dr. Robert Shapiro of New York University

Dr. Henry Stoklosa of Ciba-Geigy

Dr. Geza Szonyi of Polaroid Corporation

Drs. D. Williams and R. McGrew from Dow Chemical

Interactions with Other SUMEX-AIM Projects

The community of scientists is a valuable resource in itself. We have shared many ideas about both programming and chemistry with several other SUMEX projects. In particular:

Prof. Todd Wipke of the Chemical Synthesis Project has spent part of his sabbatical quarter at Stanford (in a DENDRAL office) and has attended several of our group meetings. We have interacted with other members of his group on common problems of graph-theory as it relates to representation and manipulation of chemical structural information.

Dr. Robert Shapiro spent a week visiting us to learn about CONGEN and related programs and their application to structures of unknown compounds obtained from reactions of various compounds with DNA nucleosides.

Bill White, a senior programmer for DENDRAL, has been working with Penny Nii, Bill vanMelle and others on programmers' aids for managing large INTERLISP programs.

The MOLGEN project has as its primary goal assisting geneticists in experiment planning. One of the new research areas of our project (see above) is also experiment planning, in our case to assist chemists in solving new structures. We have interacted with the MOLGEN project regularly to exchange ideas and discuss methods.

Although Ray Carhart has been in Edinburgh this year, we have continued close interaction with him because he is able to send and receive messages from us and run programs on SUMEX.

Critique of Resource Management

DENDRAL's View of the SUMEX Resource

Our research efforts depend heavily on adequate computer resources in order that both program development and application of the program to structural problems can be carried out. The interactive nature of program development and application demands a time-sharing environment where such interactions are possible. The SUMEX resource has been an ideal vehicle, not only in its support of time-sharing but also in the variety of languages, editors and system functions SUMEX provides which can be employed to solve special problems.

We are currently approaching the limits of our share of the SUMEX resource. The combination of new developments and our desire to offer at least trial access to a broad community of collaborators comes at a time when the demands on the resource from all projects are taxing its capabilities. We have alleviated this problem somewhat by voluntary scheduling of our development efforts, including shifting some work away from prime time. We also encourage our collaborators to avoid prime time use whenever possible, and have provided some batch capabilities to run long computations overnight. However, despite these efforts, there is no question in our minds that the pace of new developments and applications is slowed somewhat by the demands on SUMEX. We welcome any augmentation of the resource, for example, a smaller machine for applications packages, which would alleviate the current situation.

III. RESEARCH PLANS

Long Range Goals

We are developing an integrated package to aid scientists in the elucidation of molecular structure of compounds of biomedical significance. Specific steps within the current 3-year grant period are detailed in the research summary sections above.

Justification and Requirements for Continued SUMEX Use

We have successfully demonstrated to the NIH the biomedical relevance and importance of the research. The last site-visiting team was composed of about a dozen nationally known chemists and computer scientists, who were extremely thorough in their examination of the goals and methods of the project. Their endorsement is taken as strong support for the biomedical relevance of our work as well as support for the technical details of our methods.

The DENDRAL programs have been developed in INTERLISP because of the extreme flexibility of the language and the programming aids it offers. We will continue to exploit these features for trying new ideas and developing first versions of new programs, even as we map the better developed parts of our program into a more compact (and more rigid) language.

Use of Other Computational Resources

As mentioned in the research summary, we are developing a version of CONGEN for the NIH/DCRT machine and for the NIH/EPA Chemical Information System computer. We have arranged for time on these machines to test the programs. These arrangements will not affect our use of SUMEX but will ease the outside demand for access to SUMEX by chemists wanting to run CONGEN.

Recommendations for Resource Development

see Critique of Resource Management, above.

IV. FUNDING

National Institutes of Health Biotechnology Resources
Program Grant RR-00612

Principal Investigator: Carl Djerassi, Principal Investigator
Professor of Chemistry, Stanford University

Budget for Year 1 from 5/1/77 through 4/30/78
Total Direct Costs: $218,580.

3-Year Summary from 5/1/77 through 4/30/80
3-Year Total Direct Costs $698,399.

REFERENCES

[1] Bruce G. Buchanan and Dennis H. Smith, "Computer Assisted Chemical Reasoning," in E.V. Ludena, N.H. Sabelli and A.C. Wahl (eds.), Computers in Chemical Education and Research, New York: Plenum Press, 1977. P. 401

[2] Bruce G. Buchanan, "Issues of Representation in Conveying the Scope and Limitations of Intelligent Assistant Programs," in D. Michie (ed.), Machine Intelligence 9, forthcoming.

[3] Bruce G. Buchanan and Tom Mitchell. "Model-Directed Learning of Production Rules," in D.A. Waterman and F. Hayes-Roth (eds.), Pattern-Directed Inference Systems, New York: Academic Press, forthcoming.

[4] Bruce G. Buchanan and Edward A. Feigenbaum, "DENDRAL and Meta-DENDRAL: Their Applications Dimension," Artificial Intelligence, forthcoming.

[5] Raymond E. Carhart and Dennis H. Smith, "Applications of Artificial Intelligence for Chemical Inference XX. Intelligent Use of Constraints in Computer-Assisted Structure Elucidation", Computers and Chemistry, 1, 79 (1976).

[6] Raymond E. Carhart, "A Model-Based Approach to the Teletype Printing of Chemical Structures," Journal of Chemical Information and Computer Sciences, 16, 82, 1976.

[7] C.J. Cheer, D.H. Smith and C. Djerassi, and B. Tursch, J.C. Braekman and D. Daloze, "Applications of Artificial Intelligence for Chemical Inference XXI: The Computer-Assisted Identification of [+]Palustrol in the Marine Organism Cespitularia sp., aff. Subviridis", Tetrahedron, 32, 1807 (1976).

[8] C. Djerassi, R. M. K. Carlson, S. Popov and T. H. Varkony. Sterols from Marine Sources. In press.

[9] R. G. Dromey, Mark J. Stefik, Thomas C. Rindfleisch, and Alan M. Duffield, "Extraction of Mass Spectra Free of Background and Neighboring Component Contributions from Gas Chromatography/Mass Spectrometry Data," Analytical Chemistry, 48, 1368, August 1976.

[10] S. Popov, R. M. K. Carlson, A-M. Wegmann and C. Djerassi. Occurrence of 19-Nor Cholesterol and Homologs in Marine Animals. Tetrahedron Lett., 3491 (1976).

[11] S. Popov, R. M. K. Carlson, A-M. Wegmann and C. Djerassi. Minor and Trace Sterols in Marine Invertebrates. 1. General Methods of Analysis. Steroids, 28, 699 (1976).

[12] Gretchen M. Schwenzer, "Applications of Artificial Intelligence for Chemical Inference. XXVI Analysis of C-13 NMR for Mono-Hydroxy Steroids

Incorporating Geometric Distortions," Journal of Organic Chemistry, forthcoming.

[13] T.M. Mitchell and G.M. Schwenzer, "Applications of Artificial Intelligence for Chemical Inference XXV. A Computer Program for Automated Empirical 13C NMR Rule Formation," Organic Magnetic Resonance, forthcoming.

[14] Tom M. Mitchell, "Version Spaces: A Candidate Elimination Approach To Rule Learning," Proceedings of the Fifth IJCAI, 1, 305, August 1977.

[15] James G. Nourse, "Generalized Stereoisomerization Modes," Journal of the American Chemical Society, 99, 2063, 1977.

[16] Gretchen M. Schwenzer and Tom M. Mitchell, "Computer Assisted Structure Elucidation Using Automatically Acquired 13C NMR Rules," in D. Smith, (ed.), Computer Assisted Structure Elucidation, ACS Symposium Series, Vol. 54:58, 1977.

[17] D.H. Smith, (ed.), "Computer-Assisted Structure Elucidation." (ACS Symposium Series 54). Washington, D.C.: American Chemical Society, 1977.

[18] R.E. Carhart, T.H. Varkony, and D.H. Smith, "Computer Assistance for the Structural Chemist," in "Computer-Assisted Structure Elucidation," D.H. Smith, (ed.), American Chemical Society, Washington, D.C., 1977, p. 126.

[19] D.H. Smith, M. Achenbach, W.J. Yeager, P.J. Anderson, W.L. Fitch, and T.C. Rindfleisch, "Quantitative Comparison of Combined Gas Chromatographic/Mass Spectrometric Profiles of Complex Mixtures," Anal. Chem., 49, 1623 (1977).

[20] D.H. Smith and P.C. Jurs, "Prediction of 13C NMR Chemical Shifts," J. Am. Chem. Soc., submitted for publication.

[21] Dennis H. Smith and Raymond E. Carhart, "Structure Elucidation Based on Computer Analysis of High and Low Resolution Mass Spectral Data," in M.L. Gross (ed.), Proceedings of the Symposium on Chemical Applications of High Performance Spectrometry, Washington, D.C.: American Chemical Society, in press.

[22] T.H. Varkony, R.E. Carhart, and D.H. Smith, "Applications of Artificial Intelligence for Chemical Inference XXIII. Computer-Assisted Structure Elucidation. Modelling Chemical Reaction Sequences Used in Molecular Structure Problems," in W.T. Wipke, (ed.), Computer-Assisted Organic Synthesis, Washington, D.C.: American Chemical Society, 1977.

[23] Tomas H. Varkony, Raymond E. Carhart, and Dennis H. Smith, "Computer Assisted Structure Elucidation, Ranking of Candidate Structures, Based on Comparison Between Predicted and Observed Mass Spectra," in Proceedings of the Twenty-Fifth Annual Conference on Mass Spectrometry and Allied Topics, Washington, D.C., 1977.

[24] Tomas Varkony, Dennis Smith, and Carl Djerassi, "Computer-Assisted Structure Manipulation: Studies in the Biosynthesis of Natural Products," Tetrahedron, forthcoming.

[25] T.H. Varkony, R.E. Carhart, D.H. Smith, and C. Djerassi, "Computer-Assisted Simulation of Chemical Reaction Sequences. Applications to Problems of Structure Elucidation," J. Am. Chem. Soc., submitted for publication.

[26] Annemarie Wegmann, "Variations in Mass Spectral Fragmentation Produced by Active Sites in a Mass Spectrometer Source," Analytical Chemistry, forthcoming.

4.2.3 AGE PROJECT

AGE ("Attempt to Generalize")

H. Penny Nii
Edward A. Feigenbaum
Computer Science Department
Stanford University

Isolate inference, control and representation techniques from previous knowledge-based programs; reprogram them for domain independence; write a rule-based interface that will help a user understand what the package offers and how to use the modules; and make the package available to other members of the AIM community and labs doing knowledge-based systems development, and the general scientific community.

I. SUMMARY OF RESEARCH PLAN

Technical Goals

The goal of this new effort is to construct a computer program to facilitate the building of knowledge-based systems. The design and implementation of the program will be based primarily on the experiences gained in building knowledge-based systems at the Heuristic Programming Project in the last decade. The programs that have been, or are being, built are: DENDRAL, meta-DENDRAL, MYCIN, HASP, AM, and MOLGEN and CRYSALIS. Initially, the AGE program will embody methods used in these programs. However, the long-range objective is to integrate methods and techniques developed at other AI laboratories. The final product is to be a collection of building-block programs combined with a knowledge-based front-end that will assist the user in constructing knowledge-based programs. It is hoped that AGE can speed up the process of building knowledge-based programs and facilitate the dissemination of AI techniques by: (1) packaging common AI software tools so that they do not need to be reprogrammed for every problem; and (2) helping people who are not knowledge-engineering specialists to write knowledge-based programs.

Medical Relevance and Collaboration

The DENDRAL, meta-DENDRAL, MYCIN, MOLGEN, and CRYSALIS, projects have been creating intelligent agents to assist human problem solving in tasks of significance to medicine and biology (see separate sections for discussions of work and relevance). Without exception the programs were handcrafted. This process often takes many years, both for the AI scientists and for the experts in the field of collaboration.

The time has come for automating some of the activities involved in building knowledge-based programs. Close collaboration of domain experts is still necessary to provide the knowledge base, but we can reduce the system building and programming time of the AI scientists. Since we view our science as an empirical science, in which many programming experiments are conducted before we produce programs suitable for a task, reducing the programming and experimenting time would significantly reduce the time required to build knowledge-based systems.

Progress Summary

AGE SYSTEM ORGANIZATION

AGE itself is a knowledge-based system organized around five subsystems, each containing its own distinct knowledge. Most of the knowledge is about itself - what facilities it has, and how to use those facilities. Figure 1 shows the general interrelationship among these subsystems. When the system is completed,it is intended to be self-contained.

```
   SYSTEM KNOWLEDGE       SUBSYSTEMS          RESULT
                              |
     ___________         _____V______
    |  MANUAL   |....>|    TUTOR    |
    |___________|...  |______|______|
                   .         |
     ___________   .    _____V______       ____________
    |  DESIGN   |....>|   DESIGN    |....>|USER SYSTEM |
    |  SCHEMA   |...  |             |     |   DESIGN   |
    |___________| .   |______|______|     |______↑_____|
                  .          |<..................
     ___________  .     _____V______       ____________
    | BUILDING  |....>| ACQUISITION |....>|    USER    |
    |  BLOCKS   |     |             |     |   SYSTEM   |
    |___________|     |_____________|     |_____|______|
                              |<.................
                        ______V_____            .
                      |    TEST     |           .
                      |______|______|           .
                              |<.................
     ___________        ______V_____
    |EXPLANATION|....>|EXPLANATION  |
    |  SCHEMA   |     |             |
    |___________|     |_____________|
```

Figure 1 AGE System Organization

The five subsystems and their current status are described below:

1. TUTOR:

Description: TUTOR's function is to guide the user in browsing through its knowledge base called the MANUAL. The MANUAL contains a general description of the building-block components on the conceptual level, a more specific description of the implementation of these concepts within AGE, description of how these components are used and how they can be constructed by the user, and various examples.

Status: The prototype TUTOR system has been written. The MANUAL is approximately 50% complete for the current version of AGE.

2. DESIGN:

Description: The function of the DESIGN subsystem is to guide the user in designing and constructing his program (henceforth called the "object program". The necessary knowledge is represented in the DESIGN-SCHEMA. Using this schema, the DESIGN subsystem guides the user from one design decision point to another. At each decision point, the user has access to the MANUAL and to advice regarding design decision at that point. This subsystem will also keep track of what has, and what has not been specified. An appropriate ACQUISITION module can be invoked form the DESIGN subsystem so that general design and implementation specifications can be accomplished simultaneously.

Status: The DESIGN-SCHEMA is complete for the currently available building blocks. A prototype DESIGN subsystem has been written to help the user design the object program.

3. ACQUISITION:

Description: For each system component that the user must specify, there is a corresponding acquisition module that will ask the user for task specific information. The acquisition subsystem is guided by DESIGN-SCHEMA as well as more detailed knowledge about the components.

Status: About 80% of the Acquisition subsystem is complete.

4. TEST:

Description: This subsystem contains the testing and debugging facility and will rely heavily on the excellent INTERLISP debugging facility. It will be augmented by traces of the reasoning steps of the object program, and of AGE itself. Eventually the tracing capability will be extended back to the decisions made during the design phase leading to "debugging" on the conceptual level.

Status: Traces of the runs of the object program is available. However, the testing facility is minimal at the present.

5. EXPLANATION:

Description: AGE has enough information for a replay of its execution steps, and it has reasonable justifications for the actions within the building blocks. However, AGE is totally ignorant of the object task domain and has no way to conduct a dialogue about that task. In the future we hope to represent different kinds of explanation that will interface to the knowledge base of the user's domain.

Status: No work has begun on this subsystem.

BUILDING BLOCKS

Although the building-block components have roots in previous programs, they have been carefully selected and modularly programmed to be useable in new combinations. The current AGE system aims to provide the user with a subset of problem solving methods applicable to hypothesis formation [Nii:77, Engelmore:77]. This hypothesis formation framework is based on the HEARSAY MODEL [Erman:75, Lesser:77] and uses the concepts of a a"blackboard" (a globally accessible data structure) and independent sources of knowledge which cooperate to form hypotheses (see figure 2).

Within this framework, currently there are enough components to build object programs for tasks characterized by the need for:

1. integration of multiple, independent sources of knowledge,

2. multi-level representation of solution hypothesis (hierarchically organized hypothesis structure),

3. representation and manipulation of uncertain knowledge using weights assigned to inferences generated by rules,

4. problem-solving methods using goals, expectations, and events,

5. independent focus-of-attention mechanisms.

```
                      ____________
                     | FOCUSSING  |
         ...........>| (Control)  |<..........
         .           |___↑___|____|          .
         .               |   |   .           .
   ____________          |   |   .      ____________
  | HYPOTHESIS |         |   |   .     | KNOWLEDGE  |
  |(blackboard)|         |   |   .     |  SOURCES   |
  |____________|         |   |   .     |____________|
        ↑                |   V   .
        .             _____________v_
        .            |  RELEVANT     |
        .............|  KNOWLEDGE    |
                     |_______________|
```

Figure 2. Control flow of the object program - HEARSAY model

APPLICATION OF AGE

Using the current version of AGE (AGE-0), two knowledge-based program are being developed. They are:

1. AGEPUFF: a program to diagnose pulmonary function disorder. One version with approximately 50 rules has been completed. The rules used in this application are those developed for the PUFF project described elsewhere. These rules were modified to be consistent with AGE representation. The current method applied to PUFF rules is and event-oriented invocation of rules. An expectation-oriented rule processing method and MYCIN-like goal-oriented rule processing method are now being applied to the same set of rules.

2. CRYPTO: a program to solve cryptogram puzzles (a "breadboard" domain to assist program development and debugging).

These knowledge-based programs use the currently implemented problem solving framework (collection of building blocks) that was described above. That part of AGE designed to help the user construct his knowledge-based program was used to transform the PUFF rules into AGE acceptable form.

Funding Support Status

The AGE project is currently partially supported under the Heuristic Programming Project contract with the Advance Research Projects Agency of the DOD, contract number MDA 903-77-C-0322, E. A. Feigenbaum, Principle Investigator. During the next year, SUMEX core research funds will provide partial personnel support for approximately 0.9 FTE.

II. RESEARCH PLAN

Research Topics

Two specific long range research activities of the AGE effort are:

1. The isolation of techniques used in knowledge-based systems. It has always been difficult to determine if a particular problem-solving method used in a knowledge-based program is "special" to a particular domain or whether it generalizes easily to other domains. In the currently existing knowledge-based programs the domain-specific knowledge and the manipulation of such knowledge using AI techniques are often so closely coupled that it is difficult to make use of the programs for other domains. Our goal is to isolate the AI techniques that are general to determine precisely the conditions for their use.

2. Guiding users in the initial application of these techniques. Once the various techniques are isolated and programmed for use, an "intelligent front end" is needed to guide users in the application of these techniques. Initially, we assume that the user understands AI techniques and knows what he wants to do, but that he does not understand how to use the AGE program to accomplish his task. A longer-range interest involves helping the user determine what techniques are applicable to his task. That is, we assume that the user does not understand the necessary techniques of writing knowledge-based programs. Some questions to be posed are: What are the criteria for determining if a particular application is suited to a particular problem-solving framework? How does one decide the best way to represent knowledge for a given problem? There are some smaller, but by no means trivial, questions which also need answering. Is there a "best way" to write production rules which would apply to many task domains? Is there a data representation which would cover many tasks? What is the best way to handle differences in the ability of the users of the AGE program?

To correspond to the two general research goals described above, the AGE program will be developed along two separate fronts, both of which are divided into incremental development stages. The first of these fronts is the development of the ability to help build many different types of knowledge-based programs (the "generality" front). The current framework within which object programs can be developed is a variation of the HEARSAY model. The various components of this model are contained in the BUILDING BLOCK/DESIGN/ACQUISITION subsystems. The second front is the development of "intelligence" in the interaction between the user and the AGE program; i.e. moving from dialogues on "how to use the tools in AGE" to "what tools to use" (the "how-to-what" dialogue front). A solution to this problem is reflected in the TUTOR and DESIGN subsystems described above.

Research Plan

The current plan for the development of the AGE program follows:

a. Immediately, augment the current capability of AGE to build HEARSAY-like programs with a capability to build MYCIN-like goal-oriented programs. Add to the dialogue capability, an ability to discuss how to chain rules and how to specify the necessary parameters for the context-tree-like structure. b.

Within the next three years, explore other AI methods for addition to the AGE system. Begin to extract from the user some key characteristics of the task, and using that information begin to suggest appropriate knowledge representations and problem-solving techniques for the user's task. This interactive capability will be limited to match the problem-solving methods available in AGE.

c. Test the utility of the AGE system by developing a more complex application program in some task domain.

References

ENGELMORE:77 Engelmore, R.S. and Nii, H.P., "A knowledge-based system for the interpretation of protein x-ray crystallographic data," Heuristic Programming Project Memo HPP-77-2, January, 1977.

ERMAN:75 Erman, L.D. and Lesser, V.R., "A multi-level organization for problem solving using many, diverse, cooperating sources of knowledge," Proc. 4th IJCAI, 1975, pp.483-490.

LESSER:77 Lesser, V.R. and Erman, L.D., "A retrospective view of the HEARSAY-II architecture," Proc. 5th IJCAI, 1977, pp. 790-800.

NII:77 Nii, H. P. and Feigenbaum, E.A., "Rule-based understanding of signals," Conf. on Pattern-Directed Inference Systems, Hawaii, 1977.

4.2.4 HYDROID PROJECT

HYDROID - Studies in Distributed Processing and Problem Solving

Prof. Gio Wiederhold
Computer Science and Electrical Engineering
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. TECHNICAL GOALS

The objective of this research is the development of a methodology for the analysis and the implementation of distributed processing. The primary reason for interest in this area is its potential of multi-processors to break through the speed limitation barriers imposed by uniprocessing systems. Other reasons are expectations of reduced cost of computation and increase in reliability. If such a breakthrough can be achieved then the viability of the applications being developed by other projects using the SUMEX-AIM resource will be enhanced.

The rapid development of processor and communications technology has given rise to a large number of proposals for implementations of networks employing multiple processors. The computations to which these distributed systems are to be applied include heuristic decision-making problems, mathematical modelling, data reduction, and database search, as well as general purpose multi-access computing. There is however a lack of an adequate global understanding of the computational trade-offs implied by network architectures.

In order to complement the experimental results of other investigators and broaden their applicability to the system-design decision-making process, we are developing a general framework for the study of processor interaction in parameters from programs which specify the computations, rules to parameterize descriptions of networks of processors, and procedures to calculate expected system performance from these parameter sets. The framework is to be sufficiently powerful so that, when it is validated, the methods will be able to assist in the a priori assessment of the potential performance of new system alternatives or of systems with improved system components.

One of the primary tools we are using to analyze the interaction between computations and distributed processor networks is simulation. The behaviour of processor network nodes, interprocessor control and task flow, and problem decomposition all require simulation at different levels of abstraction. Analytic queuing models may provide insight into relationships in networks, but are not adequate to provide quantitative results. Simulation is not seen as the end product of the study, but as a means to develop and assess the validity of our model of the interaction of computations and processor network architecture. Where possible, mathematical results are used to assess the validity of model simulations. Most actual simulations, since they require extensive computing resources, are being done using other computers at Stanford.

A number of large computational applications are being analyzed in order to assess their potential for decomposition into modules for distributed processing. The current applications are:

a) Programs which use heuristic methods in decision-making. Heuristic programs frequently employ recursive decomposition of problems into subsidiary problems which themselves may be suitable for distributed processing.

b) Programs which use multi-faceted databases to retrieve and abstract information. The process of intelligent data retrieval and analysis often depends on data or knowledge sources which are being maintained at geographically distributed processing sites.

c) Programs which acquire data from multiple, possibly dissimilar, sensors and attempt to reduce this data to simpler hypotheses.

Further candidates for analysis are:

d) Programs which solve large numerical problems, such as those found in image processing applications.

e) Operating system functions.

How the parts of those programs work together depends on the coupling required between their processes and on the facilities to support this coupling by the hardware. Parameters which describe the computations to be simulated include:

a) The computational kernel size: the cycle and memory demand of a computational unit between interprocessor reference requirements.

b) The computation definition message size: the amount of data required to transmit sufficient information to sufficient information to initiate a computational kernel.

c) The database size: the amount of data or program text required to sustain a computational kernel, and its availability and residence in the network.

The behaviour of the system can be varied through the adjustment of other parameters. These parameters may be set to reflect the architecture of specific hardware systems, or may be varied to obtain optimum performance. In addition to obvious parameters (as the number and power of the processors), we expect the following parameter types to be important in developing an understanding of the spectrum of distributed processor architectures:

a) Interconnection density. As the density decreases, the message delay and congestion increase. This parameter will provide a high level abstraction of multi-processor connectivity schemes. Geographical distribution will increase message delay and transmission const.

b) Computational locality. A high degree of locality (of database or procedural information in the network) will enhance the probability that relevant knowledge exists in closely linked nodes, thus counteracting the effects of a low interconnection density.

c) Database viscosity. A database, including the programs required to carry out the computations at a node, may be more or less fixed to one specific node. This therefore encourages the use of certain nodes for specific functions. Many current processor networks are completely rigid in this sense, and for these networks optimal initial program and database allocations may be determined. However, we hypothesize that a greater degree of dynamic resource allocation is desirable to cope with changing loads and in order to enhance reliability. For this reason this parameter needs to be included.

d) Redundancy. In order to assess the cost and benefits in terms of responsiveness and reliability, the redundancy of database and computations will also be made a parameter. In order to utilize the redundancy well, the computational resources (programs or data) which effect system performance most must be identifiable.

e) Error rate. In order to test the effectiveness of reliability strategies, node and communications channel failures will be simulated.

An important aspect of this model is that we intend to keep the abstractions at a sufficiently high level to allow analytic and intuitive verification of the model behaviour when applied to well understood computations. In the past several large applications have been mapped into specific parallel machines, but these results are not easily transferred to new architectures. The distributed processor systems now being built may have characteristics with unpredicted effects on system behaviour. We expect to be able to use the model to find potential bottlenecks, which then will define areas where extra design attention has a high payoff.

We do not intend to build hardware which is based literally on the abstract model. We hope to verify results obtained from the model using existing distributed processor systems and, assuming that our model (with appropriate parameters describing the load and architecture) matches the given system, be able to advise on system utilization or development aspects.

We are currently active in the support of one specific architecture, the S-1 Advanced Technology Processor. This system promises to deliver very high performance through the use of quite powerful processors nodes (on the level of the DEC KL-10 or more), large shared address spaces (30 bit addresses), and a moderate number of parallel processors (16 for the Mark II version). The fact that memory is shared allows the coupling to be quite strong, but partitioning of problems with less coupling will yet incur less overhead and interference.

The hardware for the prototype processor (Mark I) node is now beginning to be operational, basic system software has been written. The PASCAL language is now available, a FORTRAN compiler is being developed, and a LISP compiler is being considered as the next task. This project is funded through Lawrence Livermore Laboratory and supported by various agencies. While this implementation is only one of the possible architectural choices for multi-processors it is encouraging that we are obtaining realistic resources to verify and test notions of parallel processing.

B. Medical Relevance and Collaboration

Many applications at SUMEX consume large quantities of computational resources. They are also being hampered by architectural limitations of the current hardware, of which the 18 bit address is probably the worst constraint. The use of multiple distributed processors may provide a means to gain the required processing capabilities in an economic manner. In this sense the medical relevance of this study is indirect. We are attempting to develop tools which will be of use in medical computation problems.

Our studies in distributed database applications have a more direct medical relevance. To this end, we are maintaining contact with Dr. Jim Fries, whose ARAMIS database network collects data for the analysis of disease progress and treatment efficacy in rheumatoid arthritis from a variety of institutions. Sharing of data to provide a broader base for analysis is also a feature of programs in cardiology and oncology in which physicians at Stanford participate. We are also discussing with Dr. Collen and his collaborators at Kaiser Permanente the benefits, costs, and trade-offs of distributed computing. The development of adequate concepts and models is important to lay a bridge which supports communications between Medical and Computer Scientists.

In each of these instances the distributed nature of the data resources leads to differences in the meaning of data items, so that simple aggregation of the data may not be valid. Distributed processing may provide a powerful alternative. We see here flow of partially processed data, i.e., information, between the nodes, rather than row data of unspecified semantic value.

C. Progress Summary

The HYDROID project has been underway since the fall of 1976. We have been involved since that time in developing a basic understanding of important problem areas in distributed processing and problem solving.

A weekly research seminar, begun in December 1976 has brought together members of the faculty and students from a variety of disciplines, and has included several speakers from application areas where distributed processing may be beneficial. Recently this seminar has been combined with a research seminar on communication issues, and this quarter a joint course on the topic of distributed computing is being presented for the first time and has attracted about 150 student from Stanford and industry. Development of industrial participation is deemed important since the resources for these systems require joint efforts.

A formalism to express the control of distributed problem solving in loosely-coupled processor networks is the subject of the nearly completed Ph.D. thesis of Reid Smith. This CONTRACT NET protocol makes the interprocessor interactions explicit. It is the cost associated with these interactions which appears to generate one of the performance boundaries for distributed processor systems.

Problems of distributed search, distributed data acquisition, and distributed databases have been simulated, as well as a model which studied network topologies.

D. Publications

1) H. Garcia-Molina, "Overview and Bibliography of Distributed Databases", Computer Science Department, Heuristic Programming Project, Stanford University.

2) H. Garcia-Molina,"Distributed Database Coupling", Heuristic Programming Project, Paper HPP-78-4, Computer Science Department, March 1978.

3) H. Garcia-Molina, "Performance Comparison of Update Algorithms for Distributed Databases",Heuristic Programming Project, Paper HPP-78-6, March 1978.

4) K. Knudsen, "Some Issues in the Design of Large Multi-Microprocessor Networks", Stanford Heuristic Programming Project, Paper HPP-77-31.

5) R. G Smith, "A Formalism for Distributed Problem Solving", Proceedings of the Fifth International Joint Conference on Artificial Intelligence, MIT AI Lab, Cambridge, August 1977.

6) R. G. Smith, "Issues in Distributed Sensor Net Design", Heuristic Programming Project, HPP-78-2, January 1978.

7) R. G. Smith and R. Davis, "Distributed Problem Solving: The Contract Net Approach", to be presented at the 2nd National CSCSI Conference, Toronto, July 1978.

8) Thomas McWilliams, Lawrence C. Widdoes, Jr., and Lowell L. Wood, "Advanced Digital Processor Technology Base Development for Navy Applications; The S-1 Project",Lawrence Livermore Laboratory, Report NCID-17705, September 1977.

9) G. Wiederhold and I. Kuhn, "Effective Services of Automated Ambulatory Medical Record Systems", Policy Analysis and Information Systems, Vol. 1 2, January 1978, pp. 21-32.

10) G. Wiederhold, "Binding in Information Processing", in preparation.

11) G. Wiederhold and R. El Masri, "A Structural Model for Database Systems", submitted for publication.

12) G. Wiederhold, "Introducing Semantic Information into The Database Schema", Proceedings of the Annual Conference of the Canadian Information Processing Society, May 1978.

13) G. Wiederhold, "The Data Model of MUMPS Globals", Proceedings of the MUMPS Users Group Meeting, June 1978.

E. Funding Support Status

The HYDROID project is currently supported under the Heuristic Programming Project contract with the Advance Research Projects Agency of the DOD, contract number MDA 903-77-C-0322, E. A. Feigenbaum, Principle Investigator.

II. INTERACTIONS WITH SUMEX-AIM

SUMEX-AIM currently provides mainly the resources for communication and documentation for the project. Recently, simulations, although initially developed at SUMEX, have been carried out on IBM 370 computers at SCIP and SLAC at Stanford due to their very high demands on computing cycles. System work for the S-1 project is carried out at the SLAC and at the SAIL facilities. Further funded work on databases will use resources at SRI. The availability of the ARPA-net has made it possible to shift work to processors where resources and funding are most appropriate. We thus enjoy a high degree of interaction with other projects involved in the problems which result from construction of large programs. Other points of contact are related to the use of the same programming languages as well as the abundance of AI expertise residing around the resource. This latter point is especially important considering that one of our aims is discovery of suitable mappings of well understood AI methods onto highly parallel asynchronous processor networks.

SUMEX-AIM is also an excellent medium for informal transmission of reports, recent results and bulletins to users with related interests and problems. The powerful screen-oriented editors available greatly enhance our capabilities for writing both text and programs.

Finally, the development of simulation programs generally requires a highly interactive computing environment - the sort of environment that SUMEX-AIM provides.

4.2.5 MOLGEN PROJECT

MOLGEN - An Experiment Planning System for Molecular Genetics

Prof. J. Lederberg (Genetics, Stanford)
Prof. N. Martin (Computer Science, U. of New Mexico)
Prof. E. Feigenbaum (Computer Science, Stanford)

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The MOLGEN project is developing a computer system capable of generating the experiment-planning sequences needed to solve given structural problems in molecular genetics. In particular we have developed a system which is capable of acquiring and representing information about genetic objects, transformations, and strategies. The knowledge base presently includes information on DNA structures, restriction enzymes, laboratory techniques, and a growing collection of genetic strategies for discovering information about various aspects of DNA molecules. Several specific subproblems such as simulating Ligase enzymes, determining safe restriction enzymes for gene excision, and inferring DNA structures from segmentation data have been explored. We have designed our effort to facilitate generalization to other domains beyond genetics in future research and applications.

The MOLGEN project has both an applications and a computer science dimension. Along the latter dimension, we seek to deepen our knowledge of the art and science of creating programs that reason with symbolic knowledge to aid human problem solvers. The task domain, molecular genetics, serves as a rich intellectual and scientific environment in which to develop and test our ideas.

The major computer science issues we are addressing during the current grant period are:

(1) Creation of a knowledge representation system with a knowledge acquisition package. The system, known as the Units Package, may be used to build a knowledge base in any suitable domain. It provides an object-centered approach for storage of both declarative and procedural information concerning all entities in the domain.

(2) Structured representation of process information. Procedures which simulate the action of the various processes in the domain form an integral part of the knowledge base. Moreover, the representation framework allows for inspection and acquisition of those procedures.

(3) Creation of program schemata and instances for general problem solving steps. Domain-independent knowledge about general problem solving methods also fits into the knowledge representation structure we have devised.

(4) Domain Specific Critics. Mechanisms for the activation of various domain specific strategies when certain predefined situations occur during the course of experiment design.

(5) Development of a specific planning strategy designed to provide high-performance for the class of genetic experiments known as discrimination experiments. The idea is based on indexing abstracted experimental designs to the types of structural features for which they have proven useful.

Along the applications dimension, we are attempting to develop tools that can benefit molecular geneticists. We believe there is substantial benefit to be derived from programs that act as "intelligent assistants" to scientists. First of all, the sheer amount of detailed knowledge a scientist is expected to know makes it likely that good experiments are being missed. Second, we believe that an intelligent planning assistant can offer some help in reasoning about the consequences of combining experimental facts in many possible ways.

A third motivation for applying artificial intelligence techniques to an experimental science like molecular genetics is to help us better understand the scientific method. The rigorous detail required for creating computer programs that assist in the performance of scientific tasks forces us to explicate concepts and procedures much more carefully than practicing scientists usually do.

B. Medical relevance and collaboration

Molecular genetics has at least two major connections to medical research. Learning about the basic mechanisms which control the operation and transmission of genetic information is necessary to understand and treat the wide range of diseases and health conditions that are genetically controlled. Also, recent developments in molecular genetics offer the promise of using genetic mechanisms to produce essentially limitless amounts of drugs ard other biomedical substances.

The MOLGEN project is a joint effort of the Computer Science Departments of Stanford and the University of New Mexico and the Genetics Department of Stanford. Major participants are Professor Nancy Martin and James Challenger of the University of New Mexico; Professor Edward Feigenbaum, Professor Bruce Buchanan, Dr. Randall Davis, Peter Friedland, and Mark Stefik of the Stanford Computer Science Department, Professor Joshua Lederberg, and Jerry Feitelson of the Stanford Genetics Department, and Professor Lawrence Kedes of the Stanford Medical School. Jim Case, a graduate student in Professor Douglas Wallace's laboratory, and Dr. John Sninsky, a molecular biologist working in Professor Stanley Cohen's laboratory, are also collaborating in the MOLGEN project.

C. Progress summary

The major effort in MOLGEN has been the creation of a knowledge management system. In addition, several specific problems which arise in genetics have been examined in sufficient detail to result in reports and/or special purpose programs. We report briefly on two such programs, SAFE written by Peter Friedland, and GA-1 written by Mark Stefik.

Knowledge management system

The success of MOLGEN as an experiment planner will depend on the quality of its knowledge base. Therefore, much of the research effort to date has been in the design and implementation of a knowledge representation and acquisition system. All of the information relevant to the planning process will be an explicit part of the knowledge base. The motivation for this aspect of the design is the necessity to expand the program capabilities in a modular fashion and to explain the rationale behind the program's planning behavior. We need to represent concepts (e.g. enzyme), instances (e.g. EcoRI), relationships among concepts, and relationships among instances. In addition, we need to represent processes. We have purposely limited the expressive power of our representations to enable us to clearly define their semantics.

The result of this work is the Unit Package. Although this package has been designed in the context of our genetics application, the package does not contain any genetics knowledge.

One important aspect of the design of the system is that the knowledge base contains knowledge about its own data representations. We have provided what we term a "bootstrap knowledge base." It contains domain independent knowledge about commonly used data types. When using our knowledge base in a new domain, an artificial intelligence researcher would probably start with the bootstrap knowledge base and then proceed to create units for the specific knowledge of his task area. Both the AGE and genetics knowledge bases have been started in this manner. The bootstrap knowledge base serves to illustrate our approach to extensibility. Most of the bootstrap knowledge base is made up of primitive datatypes. To add a new datatype to our system, one needs to provide the knowledge base with procedures for some basic operations -- such as editing and printing. Actually, the same approach is used in the unit package for defining a new datatype as is used for defining a new enzyme. The process of defining new datatypes requires, however, an understanding of Interlisp because the primitive processes in the system are grounded in that language. New datatypes must be defined together with their basic operations and entered into the knowledge base.

Knowledge base contents

The genetics knowledge base is growing rapidly. Approximately 60% of the commonly used enzymes have been characterized. A beginning has been made on the characterization of organisms such as bacteria and phages, plasmids and other vectors, and genes. Our knowledge base also contains a growing collection of genetic strategies for discovering information about various aspects of DNA molecules, as well as a hierarchy of laboratory techniques which are used to instantiate the strategies. The hierarchy of techniques includes modification, separation, visualization, sequence analysis, and bacteriological techniques at many levels of abstraction.

Safe program

The geneticist needs to predict what restriction enzymes can safely be used to excise a gene, i.e. which ones can be guaranteed not to cut the functional part of the gene. We would also like to know the approximate location of the possible cutting sites of other restriction enzymes. This would all be very easy if the complete DNA sequence of the gene was known. Sequence information is becoming more and more prevalent, but it is still uncommon to know the complete sequence of a gene. However, it is not unusual to know what protein the gene codes for and to know the amino acid sequence of a protein.

Knowing the amino acid sequence does not provide full information because of the degeneracy in the genetic code. One codon (a triplet of nucleotides) specifies only one amino acid, but up to six different codons may specify the same amino acid. The problem therefore, is combinatorically difficult. Typical proteins are up to 300 amino acids long (900 nucleotides), and all possible nucleotide sequences which would produce the protein in the three possible phases have to be considered.

The SAFE program lists the restriction enzymes that are currently stored in the knowledge base and allows the user to add new ones. Besides determining which enzymes are safe to use for gene excision in a particular DNA molecule, the program also gives the position in the amino acid sequence where the possible cutting site would be located.

GA-1 program

A common task in molecular genetics laboratories is the analysis of DNA structure from restriction enzyme segmentation data. This task is one of the simplest, although time-consuming, analysis tasks in molecular genetics. Two standard approaches to solving this problem were examined: a data-driven strategy and a model-driven strategy. These approaches are discussed and compared in terms of sensitivity to missing data, efficiency in the use of data, and other measures of performance in [Stefik, 78]. A program was designed and implemented which is superior to human performance on smaller problems, both in speed and reliability. However, on large problems human problem solvers can use extra structural constraints to out perform the program. The current program uses only constraints derived from segmentation data itself. Geneticists usually know additional information -- eg. that a given segment is on the end or that certain segments must be adjacent.

A further benefit of this work was the suggestion of two new lab techniques: combining multiple enzyme digests and incomplete digests. These ideas arose from a systematic examination of evidence and inference rules that went into building the program.

D. Publications

Challenger J., A Program for Printing DNA Structures, CIS Report 78-3 (April 1978)

Feitelson J., Stefik M.J., A Case Study of the Reasoning in a Genetics Experiment, Heuristic Programming Project Report HPP-77-18 (Working Paper) (May 1977)

Martin N., Friedland P., King J., Stefik M.J., Knowledge Base Management for Experiment Planning in Molecular Genetics, Fifth International Joint Conference on Artificial Intelligence. 882-887 (August 1977)

Stefik M., Friedland P., Machine Inference for Molecular Genetics: Methods and Applications, Proceedings of the National Computer Conference, (June 1978)

Stefik M.J., Martin N., A Review of Knowledge Based Problem Solving As a Basis for a Genetics Experiment Designing System, Stanford Computer Science Department Report STAN-CS-77-596. (March 1977)

Stefik M., Inferring DNA Structures From Segmentation Data: A Case Study, Heuristic Programming Project Report HPP-78-3 (January 1978)

E. Funding Support Status

MOLGEN is funded by two NSF grants for the 2-year period 7/1/76 through 7/1/78. These grants, including indirect costs, are:

MCS-7611935	PI:	Nancy Martin, UNM Univ. of New Mexico	Amount of $68,000
MCS-7611649	PI:	Edward Feigenbaum Joshua Lederberg, Stanford University	Amount of $110,700

Two additional grants have been requested for the period 8/1/78 through 8/1/80.

II. Interactions With The SUMEX-AIM Resource

All system development has taken place on the SUMEX-AIM facility. The facility has not only provided excellent support for our programming efforts but has served as a major communication link among members of the project. Through the SUMEX-AIM facility, program development has taken place concurrently at Stanford and New Mexico. Systems available on SUMEX-AIM such as INTERLISP, TV-EDIT, and BULLETIN BOARD have made possible the project's programming, documentation and communication efforts. The interactive environment of the facility is especially important in this type of project development.

We have taken advantage of the collective expertise on medically-oriented knowledge-based systems of the other SUMEX-AIM projects. In addition to especially close ties with other projects at Stanford, we have greatly benefitted by interaction with other projects at yearly meetings and through exchange of working papers and ideas over the system.

The combination of the excellent computing facilities and the instant communication with a large number of experts in this field has been a determining factor in the success of the MOLGEN project.

III. Research Plans

A. Project goals and plans

In exploring the three major motivations mentioned in section I.A. for creating the MOLGEN project, there are many specific subproblems. We have identified five for concentrated effort in the next two year period.

(1) Creating a more comprehensive genetics knowledge base. Expanding the knowledge base within the area of DNA structural manipulation problems.

(2) Abstracting and Saving Plans. Recognizing when newly-created experiment designs are worth saving and then generalizing those plans so they are useful for more than the specific problem environment which caused their generation.

(3) Making use of the process of hypothesis formation to help debug MOLGEN-produced experiment designs. This process is especially important in a domain like molecular genetics where incomplete knowledge about objects and processes is the rule rather than the exception.

(4) Experiment planning by analogy. MOLGEN provides an excellent environment for exploring various types of analogical reasoning. We integrate problem-solving by analogy into the experiment design system as one of the possible tools for solving subproblems.

(5) Performance evaluation as an integral part of the knowledge representation and acquisition system. We view the process of evaluating a system's performance and suggesting improvements as an AI problem solving task. The strategies for this evaluation will be stored within the same framework as all other MOLGEN planning strategies.

Each time the Heuristic Programming Project at Stanford has built another large AI program, we have learned more about how to do it better and faster next time. For example, the production rule interpreter in Heuristic Dendral (for special-purpose rules) became the general rule interpreter of MYCIN. One of the significant products of MOLGEN research will be the sets of ideas and programs for encoding and manipulating large amounts of knowledge about a scientific discipline. We have transferred some parts of the MOLGEN Units package to another project interested in building a knowledge base about AI methods and techniques. Making the tools used here available for use in new programs is an important aspect of our work and is generally important for cumulation of knowledge in the AI field. In order to do this we must reformulate the methods so they are more generally applicable and more readily combined in diverse ways.

B. Justification and requirements for continued SUMEX use.

The MOLGEN project is completely dependent on the SUMEX facility. While we have solved many of the original problems facing us in a manner useful to working geneticists, we are just in the middle phase of building a planning system. Without support from SUMEX to complete this system, many of the results of the last two years will be ineffective. In the past six months our interactions with geneticists outside of Professor Lederberg's laboratory have increased greatly. The geneticists are excited about helping us with our knowledge base. Also, with our help, they are finding useful ways to use the computing facility in their current research. Thus the serendipity of supporting MOLGEN is the creation of many useful computer research programs.

We are asked to state our requirements for continued SUMEX use. We project that our usage of processor cycles and file storage will grow to twice the current levels in the coming year.

4.2.6 MYCIN PROJECT

MYCIN - Computer-based Consultation in Clinical Therapeutics

S. N. Cohen, M.D. (Genetics and Pharmacology) and
B. G. Buchanan, Ph.D. (Computer Science)
Stanford University

I) Summary of research

The MYCIN program is an outgrowth of nearly a decade of work on DENDRAL. We are building on, and improving, many of the ideas from DENDRAL about representing large amounts of domain-specific knowledge for computer aided problem solving. The representation, use and acquisition of knowledge for computer programs has been called "knowledge engineering" [D.Michie, On Machine Intelligence, New York: Wiley, 1974]. We represent that body of knowledge as a collection of decision rules about diagnosis and therapy selection in infectious diseases. These conditional sentences are called "production rules". The production rule formalism provides an easily understood representation of facts and relations.

The MYCIN knowledge base currently consists of approximately 400 such rules. Each rule consists of a set of preconditions (called the "premise") which, if true, justifies the conclusion made in the "action" part of the rule (an example is shown below).

If 1) the gram stain of the organism is gram negative,
and 2) the morphology of the organism is rod,
and 3) the aerobicity of the organism is anaerobic,
Then there is suggestive evidence (.6) that the identity of the organism is Bacteroides.

The Medical Problem

A number of recent studies (documented in Shortliffe, 1976) indicate a major need to improve the quality of antimicrobial therapy. Almost one-half of the total cost of drugs spent in treating hospitalized patients is spent on antibiotics, and if results of a number of recent studies are to be believed, a significant part of this therapy is associated with serious misuse. Some of the inappropriate therapy involves incorrect selection of a therapeutic regimen, while another serious problem is the incorrect decision to administer any antibiotic. One recent study concluded that one out of every four people in the United States was given penicillin during a recent year, and nearly 90% of these prescriptions were unnecessary. Other studies have shown that physicians will often reach therapeutic decisions that differ significantly from the decisions that would have been suggested by experts in infectious disease therapy practicing at the same institution.

Initial culture reports from a microbiological laboratory may become available within 12 hours from the time a clinical specimen is obtained from the patient. While the information in these early reports often serves to classify the organism in general terms, it does not often permit precise identification. It may be clinically unwise to postpone therapy until such identification can be made with certainty, a process that usually requires 24 to 48 hours, or longer. Thus it is commonly necessary for the physician to estimate the range of possible infecting organisms, and to start appropriate therapy even before the laboratory is able to identify the offending organism and its antibiotic sensitivities. In this setting MYCIN plays two roles: (a) providing consultative advice that will assist the physician in making the best therapeutic decision that can be made on the basis of available information, and, (b) by its questioning of the physician, pinpointing the items of clinical data that are necessary to increase the validity of the clinical decision.

Progress Summary

In the past year our work has been guided by three fundamental objectives, roughly corresponding to use, explanation, and acquisition of medical knowledge in a symbolic reasoning program. These are discussed in the next three sections.

Use of Knowledge

A major objective of the MYCIN system has been to provide a computer-based therapeutic tool designed to be useful in both clinical and research environments. This requires development of a system that has a medically and scientifically sound knowledge base, and that displays a high level of competence in its field.

Expansion of Knowledge Base

Work on improvements to the knowledge base included both expansion into new areas and further development of existing areas of expertise. A start has been made, for instance, on rules for dealing with cystitis infections, a very common source of bacterial infection. In addition, a urinary tract infection therapy grid has been developed. This is a large table containing basic information about drug and dose selection, which functions as a basic information source for the therapy selection routines in the program.

We have also developed a facility that plots the steady state blood levels of antibiotics, based on a range of patient-specific parameters (body surface area, level of kidney function, etc.). This facility presents a very clear picture of the consequences of selecting various dose levels and intervals, and will be very useful in helping a clinician select a therapy regimen that is maximally effective without endangering the patient due to toxic levels of drug in the blood.

Extensions to the system included improving the design of rules so that they now include "justifications" and literature references. That is, rather than simply indicating that "gram negative rods in a nosocomial infection are likely to be pseudomonas", the rule now has a "justification" which is a further explanation of the reasoning behind the rule. This justification can be printed out at the users request, and may indicate, in this case, that pseudomonas is more common in the hospital setting, and hence a likely causative organism in nosocomial infections. The literature references are pointers to published articles that give further information (including original clinical studies) which serves to illustrate the reasoning behind the rule.

Finally, as a result of our recent evaluation of the system's performance on meningitis cases, we have made several improvements to the body of rules dealing with this area.

Evaluation of Competence

A formal evaluation of Mycin's performance in recommending treatment for patients with infectious meningitis was begun in May of 1977. The study design enabled unbiased comparisons to be made of Mycin's therapy recommendations with those of infectious diseases specialists and physicians of varying degrees of medical expertise. These comparisons provided a means to establish the level of competence of the Mycin system in prescribing therapy for patients with meningitis.

Ten patients with infectious meningitis of variable etiologies were selected for the study by a physician not familiar with the Mycin system. A detailed clinical summary was compiled for each patient. These summaries were presented to five faculty members in the Division of Infectious Diseases at Stanford, one senior research fellow in infectious diseases, a senior resident in medicine and a senior medical student, with a request to select therapy for each case. A Mycin consultation was performed for each patient using the same clinical summary data. As a result, a total of ten therapy recommendations, including the actual therapy the patient received, were compiled for each case.

The patient summaries were then given to eight prominent infectious diseases specialists at institutions other than Stanford. They were asked to select therapy for the patients and to evaluate the recommendations of the Stanford physicians, the student, Mycin and the actual therapy the patient received. The therapies were listed in random order for each case and the national evaluators were not informed of the identities of the ten prescribers. The evaluators used the following criteria to rate the prescribers' recommendations: "equivalent", the recommendation was identical or equivalent to his own; "close call", the recommendation was different but the evaluator was willing to rate it as an acceptable alternative; and "not acceptable", the recommendation was inappropriate.

The results of the evaluation, as summarized in the table below, demonstrate that Mycin's competence in selecting antimicrobial therapy for meningitis is comparable to that of the infectious disease faculty at Stanford.

Prescribers	Acceptability Rating
MYCIN	52
Faculty-1	50
Faculty-2	48
ID Fellow	48
Faculty-3	46
Actual therapy	46
Faculty-4	44
Resident	36
Faculty-5	34
Student	24

The acceptability rating is the cumulative total of therapy selections rated as "equivalent" or a "close call". Since there are eight national evaluators and ten cases to be evaluated a perfect score would have been 80. Mycin's therapy recommendations received the largest number of acceptable ratings.

Use of Prototypes to Guide the Reasoning

A consultation program is currently under development which uses a set of domain-specific prototypes of typical situations to guide the invocation of production rules. This program has been implemented in the domain of pulmonary physiology in which the prototypes represent the various pulmonary diseases. Each prototype consists of a group of situation-specific components, each having plausible values, likely trap values and a possible default value.

The addition of prototypes to the rule-based system allows us to:

1) Control the consultation using information given in the prototypes. The determination of values of parameters in the original consultation system (which in turn causes rules to be tried and questions to be asked) is replaced by the determination of values of components in a prototype, i.e., "filling in" the prototype. This "search" for information is more focused since it occurs within a more limited context, namely, the context defined by the prototype.

   In addition, once a prototype is chosen, a "control element" associated with the prototype is invoked. This control element explicitly states actions to be performed in "filling out" the prototype.

2) Analyze the result of the consultation when a value is unknown vs. the result when a default value is used. That is, if we provide a default value for a component, does it make a difference in the result? For example, if it is found that using a default value for a missing lab test results in a different therapy for a patient, then the user can be informed of the significance of that lab test.

3) Detect inconsistencies in the data using the Plausible Values associated with the component in the prototype.

4) Detect error conditions and suggest actions to correct them using the Likely Trap values associated with the component in the prototype.

5) Give explanations of system performance in terms of the prototypes. A trace of system actions is provided by noting which prototype is invoked at any time.

6) Classify consultations in terms of the prototypes, allowing indexing and retrieval of consultations for uses such as testing a new rule relevant to some situation.

Compilation of the Rule Base into a Decision Tree

In an effort to overcome the inherent slowness of interpreting an increasingly large set of production rules, work is underway to "compile" the rule base. The block of rules which make conclusions about a particular clinical parameter implicitly constitutes a program to conclude the parameter. Under the present control structure, that program is trivial; the rules are simply interpreted one at a time; if a predicate appears in 50 rules, it will be tested 50 times. The rule compiler, however, can construct an explicit LISP program from those rules, and, by making compile-time inferences, restructure it to eliminate redundant computation of related premise clauses. The result is an optimal decision tree, which effectively tries several rules in parallel; a single test in the compiled program (e.g. "what is the infection?") can cause a large set of rules to fail at once.

The rule compiler will thus allow the consultation program to use an efficient deductive mechanism, even as the rule base expands, while the flexible rule format is still available for explanation and debugging.

Development of a Tutoring System

The objective of this part of the project is to implement and test a computer program for tutoring medical students and physicians in infectious disease diagnosis and therapy. The problem-solving and dialogue capabilities of this tutor are distinct so that arbitrarily complex case histories may be discussed with the student or physician.

The tutor itself has the capability to discuss the problem and the expert program's solution with a student. Therefore, the teaching capability is distinct from the problem-solving expertise of the consultation program. This will enable us to study alternative teaching methods.

This work proceeds by formalizing the conventional rhetorical patterns one finds in tutoring dialogues. The resultant formalization is, as in the consultation program, independent of any particular case, so the tutor will have the capability to discuss any case presented by a student or selected from a computerized library of cases. This project is similar in style to current research in the area of "intelligent computer aided instruction." The power of these programs derives from 1) separation of problem solving ability from teaching strategies, 2) construction of a student model which describes the student's understanding with respect to the decisions of the expert program, and 3) consideration of explanation techniques that make a point clearly and capture what the student needs to know.

The initial design of the tutor has involved three major steps:

1) Modification of the control structure of the expert program to leave behind detailed traces for use by the tutorial program. These traces indicate specifically which information was necessary to apply each rule; they are significantly more complex than the traces used by the current question-answer program in MYCIN.

2) Formalization of tutorial discourse procedures. Proceeding from an extensive hand simulation, patterns in the tutorial dialogue were formalized into procedures that constitute a sequence of action options for the tutor. For example, the most complex procedure details how to discuss a topic with the student: it captures the exchange of initiative as the student requests more data and the tutor quizzes him about new conclusions he can draw. Another procedure captures the situation in which the tutor mentions rules that are related to one that the tutor and student have just discussed. These procedures have been written in stylized INTERLISP so they are translatable into an English form. This makes it easy for us to show other researchers and educators the specific rhetoric patterns, domain knowledge, and teaching strategies used by our program. There are about 20 such procedures in the current formalization of the case method dialogue.

3) Formalization of teaching strategies and the communication model. Again using stylized code, rules have been written to decide what the student knows during the dialogue, based on his behavior and past history. For example, given a student hypothesis, the tutor applies communication model rules to determine which of the expert (domain) rules have probably been considered by the student. The teaching strategies use this belief as a starting point and make an appropriate response. A range of strategies have been formulated, consisting mainly of ways to state, hint or quiz about a rule. A taxonomy of question types based on difficulty allows the tutor to choose a strategy appropriate to the student's knowledge of the (domain) rule. Thus, the communication model and teaching strategies serve to control the rhetorical options that have been specified in the discourse procedures.

4) Augmentation of the expert knowledge base. Several forms of meta-level knowledge and annotations have been added to the (domain) knowledge base. These include: a) rule schema, or templates that describe a type of rule and its import (entered by a domain expert); b) key factors, or distinguished rule clauses that are derived by the tutor from the rule schema and indicate the key piece of information being taken into account by the rule; and c) factor interrelationships, such as an indication that when one factor is known other factors are not relevant.

Explanation of Reasoning

One of the important contributions of the MYCIN work to artificial intelligence has been our persistent emphasis on an intelligent system with the ability to explain the reasoning behind its decisions. It should be able to do so in terms that suggest to the physician that the program approaches the problem in much the same way that he does. This permits the user to validate the program's reasoning, and modify (or reject) the advice if he believes that some step in the decision process is not justified. It also gives the program an inherent instructional capability that allows the physician to learn from each consultation session.

The use of a rule-based representation of knowledge makes it possible for the system to explain the basis for its recommendations. For example, if asked "How did you determine the identity of the organism?" the program answers by displaying the rules which were actually used, and explaining, if requested, how each of the premises of the rules was established. This is something which people readily understand, and it provides a far more comprehensible and acceptable explanation than would be possible if the program were to use a simple statistical approach to diagnosis.

Extension of Question Types

Additions to MYCIN's question-answering system within the last year fall into three categories: general improvements to the parsing mechanism, refinements to answers that are given, and new types of questions that can be recognized and answered.

The first major improvement to the parsing mechanism is that each word in a particular question is given a single interpretation. In particular, a word may not implicate both a context and a parameter, or more than one parameter. When dictionary pointers alone are not enough to distinguish among ambiguous interpretations of a word, the routine that generates an answer to the question will select the interpretation that best answers the question. For example, in a rule-retrieval question where a single word implicates several different parameters, the parameter that is used in the most rules will be chosen.

Another change is that it is no longer necessary to explicitly mention the name of some context in a question that is specific to the consultation. Rough semantic analysis is often successful in determining that a question is about the consultation, and in selecting the appropriate context.

A final improvement to interpreting the question is in the analysis of questions that mention more than one context from the consultation. In the past, one single context was chosen, and the selection was often arbitrary. Now, heuristics are used to choose among the contexts, and in certain questions, more than one will be used (e.g. in "DID YOU USE THE SITE OF CULTURE-1 TO DETERMINE THE IDENTITY OF ORGANISM-2?").

Questions that the system has always been able to answer now get answers that are closer to what the user wants. One example of this is that the system now differentiates between the conjunction and the disjunction of parameters in a question. In the past "HOW DO YOU USE THE FACT THAT AN ORGANISM IS A GRAM NEGATIVE ROD" would have been answered by listing the rules that use either negative gramstain or morphology rod instead of those that use both. In addition to simply listing the rules, the system now also summarizes how they are used (in the question above, it would add that most of the rules that use those two factors (gram negative, and rod) are used to conclude about the identity of the organism). When printing rules in answer to a question, the system now includes a brief explanation of when the rule will be used, and gives the user the option of seeing the rule's author, medical justification, and relevant literature references.

Some of the "holes" in the question-answerer's repertoire have been filled in. The system can now explain how it calculated the dosage of the drugs that it prescribed and it can explain how it decided to treat for particular infections and organisms.

Development of Automatic Testing System

We have added a mechanism for automatically testing the question-answering system which will be used to debug the system and to maintain it in a working state as the consultation system changes around it. Testing in the past has been tedious due to the necessary interaction with a person asking questions and verifying the validity of the answers.

The new testing program consists of two phases. The first allows a MYCIN expert to run QA and save information on sample questions that were answered correctly. The information includes the question itself, intermediate parses, and an encoding of the answer that was given. This information is stored in a file.

The second phase is undertaken after substantive changes have been made to the system. A batch job is submitted which will run a consultation, then ask the questions that were stored on the file created during the first phase of testing. This job writes a report of the comparison of the new parse and answer with what is stored in the file. The discrepancies that are noted in these reports can be used to pinpoint the sources of bugs in the QA system, if any.

Acquisition of Knowledge

A third major objective has been to provide the program with capabilities that enable augmentation or modification of the knowledge base by experts in infectious disease therapy, in order to codify knowledge in the domain, as well as to improve the validity of future consultations. The system therefore requires some capability for acquiring knowledge by interacting with experts in the field, and for incorporating this knowledge into its knowledge base.

Substantial effort was placed in updating sections of Teiresias, the knowledge acquisition system developed by R. Davis [Davis, 1976]. In particular the post-consultation review procedure was modified to utilize the existing question-answering facilities to describe how particular conclusions were made. This allows the user direct access to the full capabilities of the question-answering system while the system aids the expert in isolating the problems in the reasoning process by unfolding the reasoning tree using the methods of the QA module to describe conclusions of rules.

The actual rule-acquisition procedure has been improved by the development of a new parser to perform the actual translation of English to the internal LISP representation.

During the design and implementation of another knowledge base for an EMYCIN system, a number of issues dealing with the initial acquisition of a knowledge base were identified. They are primarily the result of having to elicit and encode substantial amounts of domain-specific knowledge from an expert who is a novice at explicating and representing this knowledge in the form of rules and parameters. We are pursuing the design of facilities which will aid the new expert during the initial stages of this knowledge explication process.

MYCIN PUBLICATIONS

Shortliffe, E H, Axline, S G, Buchanan, B G, Merigan, T C, Cohen, S N. An artificial intelligence program to advise physicians regarding antimicrobial therapy, Computers and Biomedical Research, 6:544-560 (1973).

Shortliffe, E H, Axline, S G, Buchanan, B G, Cohen, S N, Design considerations for a program to provide consultations in clinical therapeutics, Presented at San Diego Biomedical Symposium 1974 (February 6-8, 1974).

Shortliffe, E H, MYCIN: A rule-based computer program for advising physicians regarding Antimicrobial therapy selection, Thesis: Ph.D. in Medical Information Sciences, Stanford University, Stanford CA, 409 pages, October 1974. Also, Computer-Based Medical Consultations: MYCIN, American Elsevier, New York, 1976.

Shortliffe E H MYCIN: A rule-based computer program for advising physicians regarding antimicrobial therapy selection (abstract only Proceedings of the ACM National Congress (SIGBIO Session), p. 739, November 1974. Reproduced in Computing Reviews 16:331 (1975).

Shortliffe E H, Rhame F S, Axline S G, Cohen S N, Buchanan B G, Davis R, Scott A C, Chavez-Pardo R, and van Melle W J MYCIN: A computer program providing antimicrobial therapy recommendations (abstract only). Presented at the 28th Annual Meeting, Western Society For Clinical Research, Carmel, CA, 6 Feb 1975. Clin. Res. 23:107a (1975). Reproduced in Clinical Medicine, p. 34, August 1975.

Shortliffe, E H and Buchanan, B G, A Model of Inexact Reasoning in Medicine, Mathematical Biosciences 23:351-379, 1975.

Shortliffe, E H, Davis, R, Axline, S G, Buchanan, B G, Green, C C, Cohen, S N, Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system, Computers and Biomedical Research, 8:303-320 (August 1975).

Shortliffe E H, Axline S, Buchanan B G, Davis R, Cohen S, A computer-based approach to the promotion of rational clinical use of antimicrobials, International Symposium on Clinical Pharmacy and Clinical Pharmacology, Sept 1975, Boston, Mass. (invited paper)

Shortliffe E H, Judgmental knowledge as a basis for computer-assisted clinical decision making, Proceedings of the 1975 International Conference on Cybernetics and Society, pp 256-7, September 1975.

Davis R, King J J, An Overview of Production Systems, Machine Intelligence 8: Machine Representations of Knowledge (eds E W Elcock and D Michie), John Wylie, April 1977.

Davis R, Buchanan B G, Shortliffe E H, Production rules as a representation for a knowledge-based consultation system, Artificial Intelligence, Vol 8, No 1 (February 1977).

Shortliffe E H, Davis R, Some considerations for the implementation of knowledge-based expert systems, SIGART Newsletter, 55:9-12, December 1975.

Scott A C, Clancey W, Davis R, Shortliffe E H, Explanation capabilities of knowledge based production systems, American Journal of Computational Linguistics, Microfiche 62, 1977.

Wraith S, Aikins J, Buchanan B G, Clancy W, Davis R, Fagan L, Scott A C, van Melle W, Yu V, Axline S, Cohen S, Computerized consultation system for the selection of antimicrobial therapy. American Journal of Hospital Pharmacy 33:1304-1308 (December 1976).

B.G. Buchanan, R. Davis, V. Yu and S. Cohen, "Rule Based Medical Decision Making by Computer," Proceedings of MEDINFO.77, 1977).

Davis R, Applications of Meta Level Knowledge to the Construction, Maintenance, and Use of Large Knowledge Bases. Memo HPP-76-7, Stanford University, June 1976.

Davis R, Meta rules: content directed invocation, to appear, Proc ACM Conf. on AI and Programming Languages, August 1977.

Davis R, Knowledge acquisition in rule-based systems: knowledge about representations as a basis for system construction and maintenance, to appear, Proc. Conf. on Pattern-directed Inference Systems, May 1977.

Davis R, Interactive transfer of expertise: acquisition of new inference rules, to appear, Proc. Fifth IJCAI, August 1977.

Davis R, A decision support system for medical diagnosis and therapy selection, in "Data Base" (SIGBDP Newsletter), 8:58 (Winter 1977).

Davis R, Buchanan B G, Meta level knowledge: overview and applications, to appear, Proc. Fifth IJCAI, August 1977.

Clancey, W, The Structure of A Case Method Dialogue, accepted for publication in The International Journal of Man-Machine Studies, fall '78.

FUNDING

KNOWLEDGE-BASED INTELLIGENT SYSTEMS
THE MYCIN PROJECT

Bruce G. Buchanan, Principal Investigator
Adjunct Professor, Computer Science
Stanford University

National Science Foundation
Grant MCS77-02712

Budget for Year 1 from 6/1/77 through 5/30/78
Total Direct Costs $ 32,357.

2-Year Budget Summary from 6/1/77 through 5/30/79
2-Year Total Direct Costs 64,396.

II) Interactions with the SUMEX-AIM Resource

Collaborative Efforts

University of Rochester

Professor Charles Odoroff again requested access to MYCIN to use in his medical school class on medical computing. In the past, the students who examined MYCIN have provided valuable criticisms on which further developments can be based.

Carnegie-Mellon University

Dr. John McDermott, working with Professor Allan Newell, visited Stanford and has had extensive network communication with us about MYCIN's production rule formalism. He is translating MYCIN's infectious disease rules into their own OPS formalism in order to study the advantages and disadvantages of writing "fine grained"" rules, as in OPS, or "coarse grained" rules, as in MYCIN.

ISI

Drs. William Mann and James Moore of ISI visited Stanford to discuss, in part, their use of MYCIN to study man-machine communication. Because MYCIN has been developed for physicians to interact easily with the program , it was thought to be one of the most appropriate objects for their study.

EMYCIN

Part of our concern is to generalize the methods used in MYCIN, and make them available to others. We have developed a prototype of an "empty MYCIN" consultation system, called EMYCIN, that embodies most of the control structure, and none of the specific medical knowledge, of MYCIN. That is, for domains that are structured similarly to MYCIN's primary domain, the existing mechanisms for offering consultations can be coupled with knowledge (rules) of the new domain. Substantial progress has been made on providing a clear, consistent package for use by other groups.

Pulmonary Function Testing and Intensive Care Unit Monitoring

The MYCIN program and techniques have been adapted for two new medical/Artificial Intelligence applications. The first domain is the interpretation of the measurements made in the Pulmonary Function Laboratory from the Pacific Medical Center at San Francisco. This application uses the MYCIN program directly, but with a new rule set designed to diagnose the type and severity of pulmonary disease, and some measure of the reliability of the tests used in the laboratory. The second domain is the continuous interpretation problem for patients with mechanical breathing assistance in the Intensive Care Unit. This program uses the basic MYCIN architecture and concepts with extensions to handle the change in patient condition from time to time. Both of these applications are explained in more detail in the section on the PUFF project.

HEADMED

We have continued strong collaboration with Dr. J.Heiser at U.C. Irvine who has developed the HEADMED program from MYCIN. See HEADMED section for details.

Medical Consultation in Rheumatology

Preliminary design of a consultant for rheumatologic diseases was begun by Dr. R. Blum. This system would arrive at a diagnostic conceptualization of a case using a production rule type formalism as is used by MYCIN; however, fundamental revisions to the control structure would guide the rule invocation. It is anticipated that rule invocation would be in a data-driven (forward) direction rather than MYCIN's current goal-driven method. Rules would be embedded in a hierarchy of diagnostic hypotheses, conceptually similar to the INTERNIST Project at the University of Pittsburgh.

The presence of a large clinical databank would permit the addition of quantitative methods not presently utilized in the MYCIN formalism. The ARAMIS databank, a collection of clinical data on thousands of patients with rheumatologic diagnoses, will be used.

The major advantages of a consultant system compromised of both knowledge derived from clinicians and from a clinical databank would be 1) the capability for comparison of the two sources of knowledge, 2) the advantage over a pure databank approach of being able to provide advice when the data alone are insufficient, and 3) the advantage over a purely rule-based system of possessing quantitative precision when the data allow it in determining probabilities of symptom/disease association or in predicting outcome parameters.

DENDRAL

Many problems of developing complex reasoning programs in INTERLISP are common to both DENDRAL and MYCIN. We have continued close association with the DENDRAL project and are jointly working on solutions to the problems of managing large programs.

Critique of Resource Management

Management of this resource remains as professional as ever. The service from the staff, as well as from the computer, is superb.

As the computer becomes more heavily loaded we encounter more and more reluctance from physicians to use the program more than once because of the response delays. This inhibits our progress, needless to say. We are working on software modifications to alleviate the problem; we are also looking forward to acquisition of new system hardware, specifically the DEC 2020, to provide some relief.

III. Research Plans

Long Range Goals

We intend to build the MYCIN rule-based consultation program into a tool for physicians in limited areas of medicine. Dr. E.H. Shortliffe will again take over the leadership of the project and intends to become principal investigator on new grant applications that we submit.

Because of our difficulty in securing funding for a phase of the project which is neither basic research nor immediate health care delivery, we may be forced to give up the idea of moving MYCIN out of the research environment with federal support. There are many challenging research questions left to be explored both in artificial intelligence and in medical computing, while we seek a solution to the problem of funding development of health care technology, as opposed to either research or application.

Justification for Continued Use

The rule-based framework for the MYCIN consultation program has demonstrated its applicability to medical domains. In addition, other SUMEX-AIM projects depend on our continued maintenance and development of the framework (see section on Collaboration).

Because of Dr. Shortliffe's commitment, future developments of the program will be based on clinical problems which he encounters. We have not identified the precise areas at this time, however.

Computing Requirements

We perceive no additional requirements either from the system or from outside.

Recommendations

Community Development - Create a "visiting scientist" position for individuals who want to spend a few months at Stanford learning about artificial intelligence in medicine. If it carries a stipend, or partial salary support for sabbatical leave, finding qualified individuals should not be difficult.

Resource Development - Undertake, as a long range goal, solution of the problem of making large research programs available (in some form) to practicing physicians and biomedical scientists. This would require persons who understand the complexities of these programs and of the people using them, and could involve purchase and support of small computers as well.

4.2.7 PROTEIN STRUCTURE PROJECT

Protein Structure Modeling Project

Prof. E. Feigenbaum and Dr. R. Engelmore (Computer Science, Stanford)

I. Summary of research program

A. Technical goals

The goals of the protein structure modeling project are to 1) identify critical tasks in protein structure elucidation which may benefit by the application of AI problem-solving techniques, and 2) design and implement programs to perform those tasks. We have identified two principal areas which have both practical and theoretical interest to both protein crystallographers and computer scientists working in AI. The first is the problem of interpreting a three-dimensional electron density map. The second is the problem of determining a plausible structure in the absence of phase information normally inferred from experimental isomorphous replacement data. Current emphasis is on the implementation of a program for interpreting electron density (e.d.) maps.

B. Medical relevance and collaboration

The biomedical relevance of protein crystallography has been well stated in an excellent textbook on the subject (Blundell & Johnson, Protein Crystallography, Academic Press, 1976):

> "Protein Crystallography is the application of the techniques of X-ray diffraction ... to crystals of one of the most important classes of biological molecules, the proteins. ... It is known that the diverse biological functions of these complex molecules are determined by and are dependent upon their three-dimensional structure and upon the ability of these structures to respond to other molecules by changes in shape. At the present time X-ray analysis of protein crystals forms the only method by which detailed structural information (in terms of the spatial coordinates of the atoms) may be obtained. The results of these analyses have provided firm structural evidence which, together with biochemical and chemical studies, immediately suggests proposals concerning the molecular basis of biological activity."

The project is a collaboration of computer scientists at Stanford University and crystallographers at the University of California at San Diego (under the direction of Prof. Joseph Kraut) and at Oak Ridge National Laboratories (Dr. Carroll Johnson). Our principal collaborator at UCSD is Dr. Stephan Freer.

C. Progress summary

During the past year we have continued with the design and implementation of a system of programs for interpreting three-dimensional e.d. maps. Progress has been made by attacking the problem from two directions: working upward from the primary data (i.e., the array of e.d. values) to higher level symbolic abstractions, and working downward from the given amino acid sequence and other experimental information to generate candidate structures which can then be confirmed by the abstracted data.

Research which emphasized the "bottom-up" approach yielded results in several areas:

1) A new skeletonization technique, based on Greer's work, was implemented, making it considerably easier to generate an optimum skeletal representation of the map. Experience in using this method on several different e.d. maps has led to heuristics for selecting grid size, minimum density threshold, contour levels, etc. Moreover, some additional post-processing of the skeleton has been implemented to indicate connectivity, to produce a list of "model atoms" for map refinement (see below) and to provide input for graphical display.

2) A new technique for improving the quality of medium resolution (about 2.5 Angstroms) e.d. maps was formulated and tested on simple structures and on a real protein currently under investigation at UCSD. The technique utilizes the quasi-atomic structure produced by skeletonizing the map. The grid points in the map which are not removed by the skeletonization process are treated as zeroth-order model atoms. A difference map is then computed, using the difference between observed and calculated structure amplitudes, and calculated phases, for the Fourier coefficients. The difference map is then used to move the model atoms toward regions of higher density. The process is repeated until the shifts in atomic coordinates become insignificant. The standard crystallographic R value is used as one measure of quality of the map. Another measure is a subjective one of displaying the map and inspecting it in regions where the protein structure has been partially determined. This technique is currently being used as one step in the determination of the structure of cytochrome c peroxidase. The method is similar to one formulated by Agarwal in its use of a set of model atoms, whose positions are progressively refined. The use of a difference map is based on earlier methods of refinement developed by Freer et al.

3) We investigated the utility of graphics systems as a tool for acquiring model building heuristics. One system is a grey-scale display used in image understanding research at the Stanford Artificial Intelligence Laboratory. A second system is the Picture System used by Prof. Langridge's group at the UCSF Medical Center. We found the former system to be difficult to use as an interpretation aid, due to the necessity of visually synthesizing two-dimensional slices of the e.d. map. We attempted to incorporate depth into the display by making the grey-scale intensity a function of both electron density and depth, but the technique wasn't helpful. Other schemes were considered, but not implemented, due to the inordinate amount of programming effort that would be required. The Picture System at the UCSF Medical Center is similar in many respects to the system used by our UCSD collaborators, and its proximity to Stanford makes it attractive. We found the system to be a potentially valuable aid in comparing the skeletal representation of the e.d. map with the more

conventional contour map representation, as well as comparing the skeleton with known protein structures. At present, however, there is no working software for displaying a contour map and a skeletal model simultaneously, but an effort is underway to put this capability into the system.

The development of our "top-down" system (named CRYSALIS) for inferring the molecular structure from the amino acid sequence, symbolic abstractions of the e.d. map and stereochemical knowledge, is continuing. Recent additions to the system include rules for identifying main chains, side chains and bridge segments, using knowledge of expected topological properties of the skeleton and peak distributions. A new report, which focusses on the design of the CRYSALIS system, will be out shortly. The purpose of that report is to summarize the current state of the system, and to critically review it with respect to its design specifications. Although the system does make inferences from the data about some structural features of the model, it has been difficult to extend the power of the system beyond its present level. The design review was primarily motivated by a desire to see which features of the system are worth preserving and which ones need redesigning in the next version of CRYSALIS.

References:

Agarwal, R. C., and Isaacs, N. W., "Method for obtaining a high resolution protein map starting from a low resolution map", Proc. Natl. Acad. Sci. USA, Vol. 74, pp 2835-2839 (1977).

Freer, S. T., Alden, R. A., Carter, C. W. and Kraut, J., "Crystallographic Structure Refinement of Chromatium High Potential Iron Protein at Two Angstroms Resolution", J. Biol. Chem., Vol. 250,46 (1974).

Greer, J., "Three-dimensional Pattern Recognition: An Approach to Automated Interpretation of Electron Density Maps of Proteins", J. Mol. Biol., Vol. 82, pp. 279-301 (1974).

D. List of Publications

1) Robert S. Engelmore and H. Penny Nii, "A Knowledge-Based System for the Interpretation of Protein X-Ray Crystallographic Data," Heuristic Programming Project Memo HPP-77-2, January, 1977. (Alternate identification: STAN-CS-77-589)

2) E.A. Feigenbaum, R.S. Engelmore, C.K. Johnson, "A Correlation Between Crystallographic Computing and Artificial Intelligence," in Acta Crystallographica, A33:13, (1977). (Alternate identification: HPP-77-25)

E. Funding status

Grant title: The Automation of Scientific Inference: Heuristic Computing Applied to Protein Crystallography

Principal Investigator: Prof. Edward A. Feigenbaum

Funding Agency: National Science Foundation

Grant identification number: MCS 74-23461-A01

Term of award: May 1, 1977 through April 30, 1979

Amount of award: $150,200 (including indirect costs)

II. Interaction with the SUMEX-AIM resource

A. Collaborations

The protein structure modeling project has been a collaborative effort since its inception, involving co-workers at Stanford and UCSD (and, more recently, at Oak Ridge). The SUMEX facility has provided a focus for the communication of knowledge, programs and data. Without the special facilities provided by SUMEX the research would be seriously impeded. Computer networking has been especially effective in facilitating the transfer of information. For example, the more traditional computational analyses of the UCSD crystallographic data are made at the CDC 7600 facility at Berkeley. As the processed data, specifically the e.d maps and their Fourier transforms, become available, they are transferred to SUMEX via the FTP facility of the ARPA net, with a minimum of fuss. (Unfortunately, other methods of data transfer are often necessary as well -- see below.) Programs developed at SUMEX, or transferred to SUMEX from other laboratories, are shared directly among the collaborators. Indeed, with some of the programs which have originated at UCSD and elsewhere, our off-campus collaborators frequently find it easier to use the SUMEX versions because of the interactive computing environment and ease of access. Advice, progress reports, new ideas, general information, etc. are communicated via the message and/or bulletin board facilities.

B. Interaction with other SUMEX-AIM projects

Our interactions with other SUMEX-AIM projects have been mostly in the form of personal contacts. We have strong ties to the DENDRAL, Meta-DENDRAL and MOLGEN projects and keep abreast of research in those areas on a regular basis through informal discussions. The SUMEX-AIM workshops provide an excellent opportunity to survey all the projects in the community. Common research themes, e.g. knowledge-based systems, as well as alternate problem-solving methodologies were particularly valuable to share.

C. Critique of Resource services

On the whole the services provided by SUMEX have been excellent, considering the large demand on its resources. With the important exceptions of high peaks in the weekday prime-time load average, the ratio of CPU time to total wait time during program execution is usually acceptable. The facility provides a wide spectrum of computing services which are genuinely useful to our project -- message handling, file management, Interlisp, Fortran and text editors come immediately to mind. Moreover, the staff, particularly the operators, are to be commended for their willingness to help solve special problems (e.g., reading tapes) or providing extra service (e.g., and immediate retrieval of an archived file). Such cooperative behavior is rare in computer centers.

The most serious deficiency, from our point of view, is the lack of a file transfer facility between SUMEX and the computing system in the UCSD Chemistry Department. Our day-to-day collaboration with Dr. Freer at UCSD would be greatly enhanced by a reasonably fast (even 1200 baud would suffice) channel for transmitting proposed protein models, generated at SUMEX, to the Picture System at UCSD.

III. Use of SUMEX during the follow-on grant period (8/78 - 7/83)

A. Long-range goals

Our current research grant extends through April, 1979. During that time we intend to bring the structure modeling system to a level of performance that permits reliable qualitative interpretation of high resolution e.d. maps, derived from real data and a correct amino acid sequence. We also plan to exploit the flexibility of the rule-based control structure to permit investigation of alternate problem-solving strategies and modes of explanation of the program's reasoning steps. Beyond the next two years, emphasis will be placed on expanding and generalizing the system to relax the constraints of resolution and accuracy in the input data.

B. Justification for continued use of SUMEX

The biomedical relevance of the protein structure modeling project, coupled with the need for building a computational system with a significant component of symbolic inference, qualifies the project as an AIM-relevant endeavor. SUMEX provides an excellent computing environment for creating and debugging programs (in a variety of languages), for sharing and distributing information among geographically dispersed co-workers, and for keeping up with current research in other AIM areas. Our project is clearly too small to justify an independent computing facility, and other large computer centers that are conveniently accessible do not fulfill our requisites. Consequently SUMEX has been and hopefully will continue to be an integral research tool in this project.

4.3 PILOT AIM PROJECTS

The following are descriptions of the informal pilot projects currently using the AIM portion of the SUMEX-AIM resource pending funding, and full review and authorization.

4.3.1 COMMUNICATION ENHANCEMENT PROJECT

Communication Enhancement Project

John B. Eulenberg, Ph.D. and Carl V. Page, Ph.D.
Department of Computer Science
Michigan State University

I) Summary of research program.

A) Technical goals.

The major goal of this research is the design of intelligent speech prostheses for persons who experience severe communication handicaps. Essential subgoals are:

(1) Design of input devices which can be used by persons whose movement is greatly restricted.

(2) Development of software for text-to-speech production.

(3) Research in knowledge representations for syntax and semantics of spoken English in restricted real world domains.

(4) Development of micro-computer based portable speech prostheses.

B) Medical Relevance and Collaboration.

Members of our group are in touch with Dr. Kenneth Colby and his group at UCLA who are working on similar problems for a domain of people who have aphasia.

The need for such technology in the medical area is very great. Millions of people around the world lead isolated existences unable to communicate because of stroke, traumatic brain injury, cerebral palsy or other causes. The availability of inexpensive micro-processors and voice synthesizers allows development of complex experimental systems to study human communication. The knowledge gained from these experimental systems should lead in a few years to prototypes of very low cost which will permit many people to engage in the vital acts of communication required for a "normal" life in human society.

We have organized institutes to bring together the many professionals who have an interest in this area. Together with the Tufts New England Medical Center, the TRACE Center for Research and Development for the Severely Communicatively impaired of U. of Wisconsin, and the Children's Hospital at Stanford (Maurice LeBlanc). We have begun the first newsletter for dissemination in this area. Called "Communication Outlook", the first issue will be published in April, 1978. Subscribers and contributors to the Newsletter come from a vide variety of disciplines and from many countries. John B. Eulenberg helped to organize the first Federal workshop for governmental agencies who have some interest in funding work in these areas. Represented were the Bureau of Education for the Handicapped, The Veterans Administration, The Civil Service Commission, NIH, NSF, and others. We have also been in touch with United Cerebral Palsy associations at the state and national levels. Much of our effort has been in educating those medical, educational, and governmental communities with an interest in this area on the available technology since most of them are not accustomed to funding the development of high-technology systems.

C) Progress summary.

Although some facets of the research have been underway at MSU for several years, we have been using SUMEX-AIM for only about a year at this time, having received our password in March, 1977. During the last year, we have:

1) Designed and built hardware and software allowing us to transmit files to SUMEX from our Nova 2/10 at 300 baud.

2) Organized a research team of 4 students possessing background in artificial intelligence lead by Dr. Carl V. Page to start a semantics-speech generator. This group had a very primitive prototype (written in running in June, 1977. The system uses statistical, grammatical and semantic information to generate sentences by anticipation. We are organizing a similar group again this month (we have a seasonal supply AI students) to expand the semantics.

3) Converted a large program (Orthophone) for English text to speech synthesizer codes to SAIL from Algol.

4) Obtained local support for terminals and tie-lines to use the SUMEX-AIM facility. We requested these in our original proposal but were not granted them. At present, the lack of a dedicated tie-line from East Lansing to Tymshare in Ann Arbor or Detroit is a problem for us during 0600 to 0900 PST.

5) During the past year, Dr. Reid of our project refined a wheel-chair portable personal communication system for a 10 year old boy who has cerebral palsy. It is micro-computer based and can accept inputs via an adaptive switch from a series of menus displayed on a TV screen, via Morse code, or by a keyboard. Its outputs can be TV display, hard copy, spoken English, Morse code, or musical sounds. As the memory available for small systems will soon be substantial, we will need to specify the content and connection of the choice menus using the knowledge gained in our SUMEX-AIM project.

6) A Doctoral thesis in the association of knowledge sources (letter and word frequencies, syntactics, semantics, pragmatics and belief systems) for the generation of speech has been started by one of our students, Mr. James Soddy (Supervised by Carl V. Page). Mr. Soddy will use the SUMEX-AIM system during Summer 1978 to program some examples for his thesis, as a means of obtaining current information from the AI community, and to communicate with Dr. Page who will be working in Sunnyvale, Calif. for the summer.

7) We have built and tested a myoelectric interface and used it (together with a miniature FM transmitter) for input of changing muscle potentials into a computer. There is reason to believe that this means of input may provide a higher bit rate than other known means for those people who possess severe cerebral palsy.

8) We have developed software for teaching basic educational concepts to severely impaired persons. For example we have developed a "talking" system for drilling students in Bliss symbolics. Another system we have developed teaches spelling using a voice synthesizer and TV screen.

D) Up-to-date list of publications. (1976 to date)

By John B. Eulenberg

"Technical Systems Development, Headin", Interim Report, April, 1976, Experimental Applications of Two-Way Cable Delivery, NSF Grant No. APR 75-14286.

"Interactive New Hired Information Access System with Both Voice and Hard Copy Output: User's Guide to NHQUERRY", April 11, 1976 (With Steven Kludt and Jerome Jackson (Artificial Language Laboratory Report AEB 041176))

"Language Individualization in a Computer-Based Speech Prosthesis System", National Computer Conference, New York, June 9, 1976.

"Individualization in a Speech Prosthesis System", Proceedings of 1976 Conference on Systems and Devices for the Disabled, June 10, 1976.

"The LEAF Language", Interim Report, September, 1976, NSF Grant No. APR 75-14286.

"Microprocessor-Based Artificial Language for Communication Prostheses", with M. R. Rahimi, Proc. of the National Electronics Conference, Vol. XXXI, October, 1977.

"A programmable Multi-Channel Modem Output Switch", September 22, 1976, with Joseph C. Gehman and Juha Koljonen (Artificial Language Laboratory Report AEB 092276)

"SMPTE Time Code Interface and Computer-Controlled Video Switcher", with Michael Gorbutt and Dennis Phillips, Interim Report, March, 1977 NSF Grant APR 75-14286.

"Representation of Language Space in Speech Prostheses", with R. Reid and M. Rahimi, Proc. of Fourth Annual Conference on Systems and Devices for the Disabled, June, 1977.

"Administration and Management of a Computer-Based Communication Enhancement Program", with M. R. Rahimi and L. Neiswander, Proc. of Amer. Acad. for Cerebral Palsy and Developmental Medicine, October, 1977. "When [-VOICE] becomes [+VOICE]- The Phonological Competence of People Who Cannot Speak", with Carol Myers Scotton, Proceedings of the Annual Confer. of the Linguistic Soc. of America, December, 1977.

By Carl V. Page:

"Heuristics for Signature Table Analysis as a Pattern Recognition Technique", IEEE Transactions on Systems, Man and Cybernetics,Vol. SMC-7, No. 2, February 1977.

"Discriminant Grammars, an Alternative to Parsing". with Alan Filipski, Proceedings of the IEEE Workshop on Picture Processing, Computer Graphics, and Pattern Recognition, April 22, 1977.

"Pattern Recognition and Data structures". Chapter in "Data Structures in Computer Graphics and Pattern Recognition" Edited by Allen Klinger, Academic Press, 1977. "A Survey of Artificial Intelligence in Computer-Aided Instruction", with Alice Gable (Submitted to the International Journal of Man-Machine Systems, March, 1978)

E) Funding Status.

1) Current funding. Wayne County (Detroit)

Wayne Intermediate School District	$75,816	(Third year)
Northville Public School District	$41,333	(Third year)
Jackson Intermediate School District	$26,500	(Second year)
Ingham Intermediate School District	$23,700	(First Year)
Michigan State University Division of Engineering Research	$64,500	(for each of two years).
Grand Rapids Public Schools	$2,100	
Vandervoot Foundation	$5,000	

Some of this money has been used to purchase equipment which is the property of WCISD for use in a demonstration classroom in an elementary school. Commitments in the grants have prevented us from using very much of these funds to support long range goals such as those communicated to SUMEX-AIM. However, the special communication devices, student and other research facilities provide the critical mass which will allow us to do the work that we have proposed. The main value of SUMEX-AIM to us is to allow experimentation with AI technology in order to develop the theory to develop intelligent speech prostheses.

2) Pending applications and renewals.

Oakland County Intermediate School District (Application being considered after negotiation)	$100,000.
Genessee County Intermediate School District	$100,000.
Tuscola County Intermediate School District	$20,000.
Livingston County Intermediate School District	$50,000.

As one can see from this list of sources, there is a lot of interest in this area from agencies which are not experienced in funding high-technology and research.

II) Interactions with the SUMEX-AIM resource.

A) Examples of medical collaboration and medical use of programs via SUMEX.

The faculty in the MSU College of Human Medicine who teach medical decision making were shown a demonstration of the SUMEX system, MYCIN and PARRY. We have agreed to present a demonstration of PUFF to Dr. Clyde Flory, an allergist who is the most knowledgeable person in our area in pulmonary studies. We intend to explore the possibility of a table-driven program for the treatment of allergy with Dr. Flory. If we decide to undertake the development of this, we will send another proposal to SUMEX-AIM. A member of our Medical School faculty, Dr. Richard Ropple, an expert on myoelectronics, is a member of of our research group.

The Dean of our College of Human medicine visited our laboratory in April, 1977 and we have been asked to apply for inclusion in the University's Clinical Center as part of the Rehabilitation Medicine Program.

B) Examples of sharing, contacts, and cross-fertilization with other SUMEX-AIM projects.

1. We have met with Dr. Kenneth Colby on many occasions including the SUMEX-AIM workshop in June, 1976. Our work in many ways complements his and we have had several worthwhile interchanges of information. We are converting our major software for speech generation and adaptive inputs to the SUMEX-AIM system in part so that they can be used by Dr. Colby and his group.

2. Mr. Douglas Appelt, a doctoral student at SU-AI was our principal systems programmer last summer. He is currently doing research in the same area as ours with DR. Gary Hendrix of SRI. We have used his knowledge of your system (via the message sending routines) to assist us in starting our project.

C) Critique of resource services.

Our use of SUMEX-AIM has been seasonal with most programs run during Spring Term. We have used your system for work that could not be done conveniently on our computers. We have been pleased with the system and find it as easy to use as one that is close to us geographically. Dr. Page will be working in the Bay Area this summer and plans to visit your facility as well as use it to keep in touch with the work at East Lansing.

III) Follow-on SUMEX grant period (8/78-7/83).

A) Long-range user project goals and plans.

We want to do fundamental research in artificial intelligence in the context of the generation of speech from very minimal amounts of input. This problem seems closely related to the understanding of speech. It seems that the methods of representation of knowledge used for speech or vision understanding can be used in a natural way for fluent generation of speech. Our area seems almost unique in AI in that it is socially desirable (without question). Even relatively primitive systems can improve the quality of life for hundreds of thousands of people.

Major long range goals are:

1) To do research in transposing the vocal tract to another region of the body in which an individual has suitable myoelectric control for the generation of speech.

2) To define a suitable system of semantics and to encode world knowledge in that system that would be useful for the generation of speech fluently.

3) To discover primitive operations on semantics which allow new and appropriate combinations of speech to be generated. (Using other sources of knowledge.)

4) To develop means for a severely handicapped individual to program and personalize his or her own speech and environmental control system.

5) To study means of using speech output to aid blind persons both through experiments with simplified text to speech devices and through means of teaching blind persons to write in cursive.

6) To study the effect of communication aids technology on the psychological assessment of individuals previously thought to be retarded and to study the consequences of this technology for the educational system.

7) To improve the prosodic qualities of generated speech, using its semantic aspects.

8) To design portable speech prostheses which allow maximum use of state of the art knowledge in speech generation.

9) To develop an experimental base for studying how the concepts which are articulated in speech are manipulated by individuals at differing states of mental organization

10) To study the potential for speech generation systems as a means of stimulating autistic children.

11) To develop voice recognition systems which will aid individuals with limited speech to develop their full potential.

(We don't expect to finish all of these by 1983.)

B) Justification for continued use of SUMEX by your project.

1) We need to use many sources of knowledge represented in computers to do our work, similar to many SUMEX users.

2) We know kindred spirits in the AI community who possess goals similar to our long range goals.

3) We have substantial hardware and software expertise which we are willing to share.

4) The payoff to society of our research in terms of the improved quality of millions of human lives seems great.

5) This area does not have a traditional means of support for research separate from development which makes your support vital for our long range goals. 6) Our area is very interdisciplinary and the communication aspects of SUMEX-AIM will be increasing valuable to us.

C) Plans For Other Computational Resources.

We have available to us three mini-computers: a Nova 3/10, a PDP-11/34 and a PDP-11/45 as well as the CDC 6500 and CDC 6400 of our central computer facility. Our demonstration classroom in a Detroit suburb will open soon making a Nova available to students who experience severe communication and motoric handicaps. None of our small systems possess AI software. We hope to develop prototype systems on SUMEX-AIM within the next year the can be used in our demonstration classroom. We believe that the injection of quite small amounts of AI technology into the speech generation systems can produce significant improvements in the communication and educational processes. We will be in a position to measure the effectiveness of the AI tools which we try. If we can make a case on a cost effectiveness basis, there are sources of support to acquire appropriate hardware to service a larger group of students and also support our research. At this point we do not feel the need to try to acquire a PDP-10 or other machine suited to AI research because although important to us, the number of machine cycles we require is relatively small.

D) Recommendations for Future Community and Resource Development.

1) We would find it helpful if there were means within SUMEX-AIM to assist in the transfer of prototypes written for your system to various common minicomputers.

2) We have available software to assist handicapped individuals in using computer systems. We would be happy to facilitate the use of SUMEX-AIM resources by a blind programmer by modifying appropriate software.

4.3.2 COMPUTERIZED PSYCHOPHARMACOLOGY ADVISOR

A Computerized Psychopharmacology Advisor

Jon F. Heiser, M.D.
Dept. of Psychiatry and Human Behavior
University of California at Irvine

I. Summary Research Program

A. Technical Goals

We propose to develop a computer-based automated system for education and consultation in clinical psychopharmacology. Our technical goals are envisioned in three phases:

- To develop a model of expert teaching, consulting and decision-making in clinical psychopharmacology.
- To implement this model on a computer system which responds in real time and communicates in natural language.
- To evaluate the performance of the system as a teaching and consulting aid.

B. Medical Relevance and Collaboration.

1. Medical Relevance.

For many years, it has been recognized that potent psychopharmacological agents are frequently used in an unsystematic manner. There are at least 50 discrete syndromes currently identified in clinical psychiatry which have unique hierarchies of plausible pharmacological treatments. Each therapeutic regimen in each hierarchy may involve several classes of drugs which can be preferentially ranked. A particular member of a class of drugs may occasionally be recommended on the basis of a patient's medical history, family history, response to previous treatments, current physical status, or current mental status. In addition, each treatment program has its own set of potential side effects, adverse reactions and drug-drug, drug-host, drug-age, drug-gender, drug-state-of-health, and drug-other treatment interactions.

Conventional sources of information for education or verification (books, journals, lectures, and seminars) are seldom quickly accessible or specifically pertinent. A traditional alternative is to consult a specialist. In addition to availability, reliability and validity, a good consultant has the ability to understand questions in their proper context and sequence, to give advice which can be explained or documented as needed, and to provide follow-up consultations which incorporate new information from clinical developments or additional expertise.

Our research on the Clinical Psychopharmacology Advisor is directed towards implementing all of the characteristics of a good consultant, which have only been outlined above, in a functional computer program. To our knowledge, no other computer program currently available or under development fulfills all of these goals in clinical psychopharmacology.

2. Collaboration.

a. Principal Investigator: Jon F. Heiser, M.D., Associate Clinical Professor, Department of Psychiatry and Human Behavior, University of California, Irvine.

b. Co-principal Investigator: Ruven E. Brooks, Ph.D., Assistant Professor, Department of Information and Computer Science (ICS) and Department of Psychiatry and Human Behavior, University of California, Irvine.

c. Resident Physicians, Department of Psychiatry and Human Behavior, University of California, Irvine:

Bronco R. Radisavljevic, M.D.	(March 1977)
Steven J. Smith, M.D.	(June 1977)
Frank S. Floca, M.D.	(January-March, 1978)
Neal R. Cutler, M.D.	(April-June, 1978)

d. Research Associate: Joan P. Ballard, Ph.D.

e. Medical Students, California College of Medicine, University of California, Irvine:

Clifford Risk	(October-December, 1975)
Dana W. Ludwig	(October-December, 1975)
Sue A. Clear	(May-September, 1976)
Neil R. Shocket	(April 1977-present)

f. Pharmacists, University of California Irvine Medical Center:

Pierre J. Menard, Pharm. D.	(January-June 1977)
Michael S. Toole, Pharm. D.	(September 1976-present)

g. Undergraduate Computer Science Majors, Department of Information and Computer Science, University of California, Irvine.

Darryl Hansen	(March-June, 1977)
Thomas E. Holthus	(March 1977-present)

h. Local Advisory Panel, Department of Psychiatry and Human Behavior, University of California, Irvine:

Louis A. Gottschalk, M.D.
Edward Kaufman, M.D.
John Kramer, M.D.
David Shore, M.D.

i. National Advisory Panel:

Ross J. Baldessarini, M.D.
Neuropharmacology Laboratory
McLean Hospital
Brookline, Massachusetts 02168

John M. Davis, M.D.
Illinois State Psychiatric Institute
1601 West Taylor Street
Chicago, Illinois 60612

Max Fink, M.D.
Department of Psychiatry
State University of New York at Stony Brook
Stony Brook, Long Island, New York 11794

Frederick Goodwin, M.D. (Tentative)
National Institute of Mental Health
9000 Rockville Pike, Building 10, Room 3S205
Bethesda, Maryland 20014

John H. Greist, M.D.
Department of Psychiatry
University of Wisconsin
Madison, Wisconsin 53706

Leo E. Hollister, M.D.
Departments of Medicine and Psychiatry, Stanford University
Veterans Administration Hospital
Palo Alto, California 94302

James W. Jefferson, M.D.
Department of Psychiatry
University of Wisconsin
Madison, Wisconsin 53706

Donald F. Klein, M.D.
New York Psychiatric Institute
722 West 168th Street
New York, New York 10032

George M. Simpson, M.D.
Psychopharmacology Unit, University of Southern California
Metropolitan State Hospital
Norwalk, California 90650

Robert L. Spitzer, M.D.
New York Psychiatric Institute
722 West 168th Street
New York, New York 10032

Zebulon C. Taintor, M.D.
Rockland Research Institute
Rockland State Hospital
Orangeburg, New York 10962

C. Progress Summary.

After carefully reviewing alternative contexts in which to generate the Psychopharmacology Advisor, we selected the EMYCIN software. MYCIN is a computer program which functions as a consultant in the diagnosis and treatment of infectious diseases. EMYCIN is a version of the MYCIN system with all knowledge and reference to infectious disease removed but with all features for diagnosis, treatment recommendation, explanation and knowledge acquisition retained. Our choice of EMYCIN was determined primarily by the availability of the EMYCIN software (including a significant amount of professional consultation, collaboration and system maintenance supplied by the MYCIN staff), the suitability of EMYCIN to most of our initial design considerations, and the desire of the MYCIN staff to test the EMYCIN software in a different clinical domain. EMYCIN has thus become our working model of expert teaching, consulting and decision-making in clinical psychopharmacology.

Our initial goal has been to develop a small but fully functioning Clinical Psychopharmacology Advisor. Approximately 250 rules, utilizing about 120 clinical parameters, have been developed and are used to diagnose and recommend therapy. The system, informally called HEADMED, currently has sound knowledge about the differential diagnosis of the major affective disorders, schizophrenia and the general category of organic brain disorders. The Psychopharmacology Advisor has perfunctory information concerning paranoid disorders and personality disorders. The system has a rudimentary knowledge of the Minnesota Multiphasic Personality Inventory (MMPI). HEADMED presently has skeletal knowledge about neuroses, behavior disorders and substance abuse and about organic brain disorders, regarding both type of brain disorder (e.g. delirium or dementia) and cause of brain disorders (e.g. intoxication or trauma. the program knows nothing about child psychiatry, sexual disorders and other psychiatric conditions.

The HEADMED software currently has the capability of recommending a drug treatment, if indicated, and of cautioning about potentially harmful interactions with a compromised host and with other chemical substances. The system also can print out advice concerning dosage and duration of therapy, pharmacokinetics, and warnings about common side effects and possible adverse reactions. We are beginning to develop data structures which can utilize this knowledge to compute diagnostic formulations and therapeutic plans which are highly specific to the unique properties and circumstances of a particular patient.

Work in progress is also concerned with improving the data input characteristics of the system. For example, it is common to describe a patient's thought pattern as "loose". A good tutor or consultant would immediately ask how loose: at the "paragraph" level (hardly a disorder as anyone composing a manuscript can certify), at the sentence level, the clause level, or within a word (possibly producing neologisms)? Furthermore, an experienced teacher might ask a novice student to discriminate between a loose association which is creative, a loose association which is humorous, and a loose association which is pathological.

As an initial step towards describing clinical phenomena reliably and validly, we are greatly expanding HEADMED's synonym list by interviewing various clinicians, abstracting a large number of psychiatric hospital charts, and studying results of computer consultations with HEADMED both for unanticipated responses and accepted but misinterpreted responses. Within the system, we are making extensive use of EMYCIN features called SPECIAL and EXTRASPECIAL to better anticipate how various users might describe patients and to quiz them immediately when their responses are ambiguous or unanticipated.

Another project consists of continuing to expand and modify the rules for diagnosis and treatment recommendation. Currently we are experimenting with rules which have more elaborate descriptions of the clinical disorders in their premise or left-hand side. We are just beginning an extensive elaboration of the pharmacological content of the system.

Work has also begun on modifying the EMYCIN therapy module, our first attempt at actually changing the basic form of the software. Presently, EMYCIN selects therapies by multiplying the certainty factor (CF) of each diagnosis by the CF of every treatment for that diagnosis. In most cases of psychiatric interest, there are several probable diagnoses and relatively few classes of psychopharmacological medications. It is common, therefore, for the same treatment to be recommended several times. Using the standard EMYCIN therapy algorithm, too many treatments acquire high CF's. Not uncommonly, therapeutic recommendations for subordinate diagnoses attain higher CF's than a less common but more specific and presumably better treatment recommendation for the primary diagnosis.

We have modified the EMYCIN software so that only the top diagnosis is used in choosing a therapy. Later, we plan to give the user various options such as treatment for top diagnosis only, treatment for any other diagnosis alone, or "broad-spectrum" treatment for any combination of the diagnoses (listed or unlisted) in any hierarchical order (for example, "I don't like diagnosis 2 at all, so throw that out. Now what would be your recommendation if the remaining diagnoses were inverted in rank? Suppose that the primary diagnosis were Mania, which you didn't even consider; how would that affect your recommendations?")

Demonstrations of the Psychopharmacology Advisor have been given to numerous groups and individuals on an informal basis. The following formal demonstrations of HEADMED were given in the past year:

1. VI World Congress of Psychiatry, Honolulu, Hawaii, 29 Augsut 1977.

2. Department of Computer Science, University of Hawaii, Honolulu, Hawaii, 01 September 1977.

3. Department of Psychiatry, University of Wisconsin, 21 September 1977.

4. Richard F. White, M.D., Director, Squibb Professional Services, 11 October 1977.

5. George M. Simpson, M.D., Professor of Psychiatry, University of Southern California, Metropolitan State Hospital, Norwalk, California, 03 November 1977.

6. Rockland Research Institute, Orangeburg, New York, 05 January 1978.

7. New York Psychiatric Institute, New York, New York, 06 January 1978.

8. NIMH Site Visiting Team, Extramural Psychopharmacology Research Division, 23 January 1978.

9. Department of Psychiatry, University of Texas Medical Branch, Galveston, Texas, 02 February 1978.

10. American Psychiatric Association Annual Meeting, Atlanta Georgia, 10 May 1978.

Many useful suggestions and critical insights are generated by these demonstrations.

D. List of Relevant Publications.

1. Heiser, J.F., Colby, K.M., Faught, W.S. and Parkison, R.C. Can Psychiatrists Distinguish a Computer Simulation of Paranoia from the Real Thing? Submitted for publication.

2. Heiser, J.F., Brooks, R.E. and Ballard, J.P. Artificial Intelligence in Psychopharmacology. Abstracts - VI World Congress of Psychiatry, Honolulu, Hawaii, 28 August - 03 September 1977, page 135.

3. Heiser, J.F., Brooks, R.E., and Ballard, J.P. A Computerized Psychopharmacology Advisor. Scientific Proceedings in Summary Form, The 131st Annual Meeting of the American Psychiatric Association, Atlanta, Georgia, May 1978 (in press).

4, Heiser, J.F., Brooks, R.E., and Ballard, J.P. Progress Report: A Computerized Psychopharmacology Advisor. Proceedings of the 11th Collegium Internationale Neuro-Psychopharmacologicum, Vienna, Austria, July 1978 (in press).

5. Verbal Reports:

   a. Heiser, J.F. Computer-Aided Diagnosis of Psychiatric Patients. Presented to the Research Meeting, School of Engineering, University of California, Irvine, 07 October 1976.

   b. Brooks, R.E. and Heiser, J.F. An application of Artificial Intelligence to Psychiatry. Presented to:

      Indian Institute of Technology, Madras, India, 28 September 1976, and

      Madras Christian College, Madras, India, 03 October 1976.

E. Funding Support Status.

1. The Principal Investigator, Co-Principal Investigator, Resident Physicians and Pharmacists have all been full-time employees of the University of California, Irvine or the Veterans Administration, and have participated in this research as part of their assigned duties or in their spare time.

2. Additional support in the form of Office and Laboratory Space, Clerical Assistance, Peripheral Data Processing Equipment, Supplies and Expenses for Traveling to Professional Meetings has also been provided by the University of California, Irvine.

3. The Medical Students have worked on this project either during elective periods for academic credit or during free time with support from NIMH Undergraduate Training Program Grants:

   Title: A Computerized Psychopharmacology Advisor
   Principal Investigator: Jon F. Heiser, M.D.
   Funding Agency: National Institute of Mental Health (NIMH)
   Undergraduate Training Program
   Total Award: $2400
   (2 $600 awards to Sue Arrigo Clear and 2 $600 awards to Neil R. Shocket)
   Date: 1976-present

4. Undergraduate Student Darryl Hansen worked on this project during UCI ICS Department Course 199, Independent Study. Undergraduate Student Thomas E. Holthus worked on this project during UCI ICS Department Course 199, Artificial Intelligence Project. Mr. Holthus also spent the Summer Quarter 1977 working on the project with support from the National Science Foundation (NSF) through an Undergraduate Research Training Grant to the Department of Information and Computer Science.

5. Since October 1977, Mr. Holthus has been employed half-time by the University of California, Irvine as a Research Technician in the Department of Psychiatry and Human Behavior. Mr. Holthus is assigned to work on this project and is paid $3.67 per hour.

6. A modest amount of additional support has been obtained from:

   Title: A Computerized Psychopharmacology Advisor
   Principal Investigator: Jon F. Heiser, M.D.
   Funding Agency: Anne R. Issler Endowment Fund
   Department of Psychiatry and Human Behavior
   University of California, Irvine
   Total Award: $552.50
   Dates: Awarded 06 January 1978 for an indefinite time period

7. A grant application was submitted to the National Institute of Mental Health on 01 November 1977. A grant application was submitted to the Veterans Administration on 01 January 1978. Copies of these grant proposals are available from the Principal Investigator.

8. The Director of Professional Services, E.R. Squibb and Sons Pharmaceutical Company, has offered to support Professional Collaboration through Squibb's panel of distinguished consultants.

II. Interactions with the SUMEX-AIM Resource.

A. Collaborations and Medical Use of Programs via SUMEX.

1. The MYCIN group has collaborated with our group since work on the Psychopharmacology Advisor began. The MYCIN group supplies invaluable software support to the EMYCIN program. Our group has participated in writing documentation of the EMYCIN software which presumably will be useful to all EMYCIN users. In addition to discussions of various issues by LINKs, MESSAGEs, phone calls and attendance at workshops, Dr. Heiser and Dr. Brooks have made several trips to the MYCIN headquarters, and Mr. Fagan of the MYCIN group has visited the HEADMED team.

B. Sharing and Interactions with Other SUMEX-AIM Projects.

1. Collaboration with Kenneth Mark Colby, M.D. and members of the Higher Mental Functions Project begun last year has continued in the form of writing and submitting for publication a paper reporting a "Turing Test" which was performed on-line on SUMEX, with the psychiatrist-judges located at the University of California, Irvine, the patient-person at the University of California, Los Angeles (UCLA) and PARRY at SUMEX. Prepublication copies of this paper (see I.D.1. above: Heiser et al. Can Psychiatrists Distinguish a Computer Simulation of Paranoia from the Real Thing?) are available upon request. In addition, demonstrations of the PARRY and DOCTOR programs have been given on-line, using SUMEX, to various groups of mental health professionals, computer scientists and other qualified and interested individuals.

2. The Principal Investigator visited Dr. Eulenberg and Dr. Page of the Communication Enhancement Project at Michigan State University on 23 September 1977.

3. As noted last year, the Principal Investigator first became acquainted with the SUMEX-AIM resource and initiated his collaboration with the MYCIN project by attending the first SUMEX-AIM Workshop at Rutgers University in June 1975. Dr. Heiser has continued to benefit from these workshops.

C. Critique of Resource Management.

In general, we find the SUMEX resource a hospitable environment. Despite periods of heavy load, we find that sufficient resources are available when we need them. The operating system and associated utility subsystems are flexible and powerful, and the mail system is uniquely valuable. Finally, we consider the practice of keeping system documentation on-line to be vital to a remote user community.

Our main criticism of the SUMEX resource regards the documentation for the EMYCIN software. As mentioned, we have frequently found need to contact the MYCIN staff with questions about the software. While they have always been generous with their time and good humored in answering our questions, most of these contacts would have been unnecessary if adequate documentation existed. Current industry practice is to generate about two pages of documentation per page of code listing. The quantity of documentation for MYCIN and EMYCIN combined falls well below this amount. We note, in this regard, that modern programming practice dictates the creation of most of the documentation before any of the code is written.

III. Research Plans (August 1978 - July 1981)

A. Long Range Project Goals and Plans.

1. Evaluation of the Psychopharmacology Advisor.

When the performance of the Psychopharmacology Advisor approaches an optimal level in the judgment of the Principal Investigators and the Advisory Panels, a formal evaluation will be performed. Elaborate plans have been made for three types of evaluation: as a simulation of the Principal Investigator; as a national expert; and as an actual psychopharmacology advisor. In each evaluation the system will be tested on two sets of cases: one which represents the population of patients likely to be encountered in practice, thereby measuring whether HEADMED can do well what it must do most often; and one which represents unusual or exceedingly complicated cases, thereby measuring whether the program can do well in situations where usual practices may not suffice. Details of the evaluation plans are available upon request.

We will also evaluate the EMYCIN formalism, regarding both its inherent properties as a consulting algorithm and its appropriateness for the domain of clinical psychopharmacology. If the results of the evaluation of either EMYCIN's formal properties or EMYCIN's applicability to the clinical psychopharmacological domain are less than ideal, which is almost certain to be the case, then we plan to construct a new system, HEADMED-II, incorporating what we have learned into hardware and software technologies then available.

B. Justification and Requirements for Continued SUMEX Use.

As mentioned in the preceding paragraph, we consider use of the EMYCIN software as integral to our project, at least, for the next two years. While we previously contemplated adapting the EMYCIN software to our local DECsystem-10 environment, we have since discovered that the conversion process would be tantamount to writing a new system. Therefore, at least until we have learned enough about the domain of clinical psychopharmacology to know how to supersede the EMYCIN formalism, we will have a continuing need for the SUMEX resource.

Our typical work pattern over the past year has been to hold intense working sessions of approximately a month followed by 2 to 3 month periods of only occasional system use. During our working sessions, our share of processor time has been adequate. During the coming year, we expect our working sessions to occur about every other month, resulting in, approximately, a doubling of our need for processor time.

At least part of this increased need for processor time can be reduced by increased disk space. Currently, the 300 page allocation on our main account is not sufficient to hold a SAV file for a loaded system; as a consequence, we use substantial CPU resources in repeatedly reloading our system. An increase in disk allocation to 1000 blocks would reduce this use.

Another need we have for increased resources results from the large number of short-term staff on our project. Since we have residents and medical students for only 2 or 3 month periods, we need to get them up to speed as quickly as possible. The ability to establish our own on-line bulletin board, with sufficient disk space (150 pages) to support it, would be useful in this regard.

C. Our Needs and Plans for Other Computational Resources, beyond SUMEX/AIM.

Our only immediate need for other computational resources beyond SUMEX/AIM is for local, high-speed printing, preferably combined with local file storage. Our current slow-speed printing is unsuitable for listings of large rule sets or of system code. The planned acquisition of a 1200 baud printing terminal may substantially reduce the problem.

Our future plans will depend greatly on the outcome of our current effort. If the EMYCIN formalism proves suitable for our domain, we may find the conversion effort sufficiently worthwhile to transport EMYCIN to our local environment. If we discover that a major redesign is needed, we will make our future computing plans in light of that design.

D. Recommendations for Future Community and Resource Development.

We suspect that the documentation problem exemplified in the EMYCIN software also exists in other SUMEX projects. If software developed for these projects is usable only by the original authors, then much of the impact of the work will be lost. We therefore suggest an increased emphasis on system design and documentation tools. Something along the line of the Programmer's Workbench concept, which links together global design, local design, and code, would probably be useful.

4.3.3 ORGAN CULTURE PROJECT

Application of Computer Science to Organ Culture

Professor Robert K. Lindsay and Dr. Maija Kibens
The University of Michigan, Ann Arbor

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals

The goal of this research project is to develop, using artificial intelligence techniques, a programmed model of organ culture that will assist histologists in the design and interpretation of their experiments.

In organ cultivation, one strives to create an environment outside of a living organism that will preserve the physiological and structural relationships among tissues of an organ fragment (explant). Culturing organ explants is an important technique in biomedical research on disease processes, nutrition, physiology, and medicinal effects. Experiments performed with organ explants have the advantages of being safer, quicker, less costly, and more easily interpreted than in vivo experiments. In contrast with cell culture, organ culture provides a more realistic environment for studying processes in healthy and diseased tissues and can provide information on the interactions between tissue types. It has been shown to be particularly helpful in studying diseases that are specific to one organ or that manifest themselves differently in different organs, such as cancer.

Organ cultivation involves techniques for obtaining and preparing explants, creating media, and assessing the condition of the explants at various time intervals. The first of these techniques involves primarily adeptness at physical and manipulative skills and is outside the scope of this project. We would, however, like to develop a computer model of the cultivation and evaluation processes. This model would be able to represent the knowledge which histologists have concerning the biochemical processes of life, the optimal structure of cells within an organ, the components of different culture media, staining and fixation techniques for microscope viewing, and the format of experimental designs. The desired computer system should be able to interpret microscope pictures of organ explants and store a representation of their appearance. It should also be able to interpret natural language descriptions of organ explants and store this information in the same internal format as that for images. The system should be able to access this knowledge base to answer a researcher's inquiries about particular explants and make comparisons between explants. It should also be able to interpret the relative success of different cultivation techniques and guide a researcher in the planning of future experiments.

B. Medical Relevance and Collaboration

Our research is being done in collaboration with Professors Raymond Kahn, Theodore Fischer, and William Burkel of the Department of Anatomy, the University of Michigan Medical School. During the three years that we have been in contact, they have been working to define the factors that are necessary for the long-term (9 days or longer) maintenance in vitro of prostate and lung explants. The results of their efforts will enable other researchers to maintain organs while studying the etiology of various diseases affecting these organs or to test the direct action of drugs and hormones. Separate procedures must be developed for each organ type since the maintenance requirements are different for each organ. The cultivation of prostate explants was funded by an NIH contract. Such a technique would be useful for studying the etiology of benign hyperplasia and carcinoma of the human prostate. The cultivation of lung is funded by an NIH grant and is done in collaboration with Dr. Paul Weinholdt, a biochemist at the Veterans Administration Hospital in Ann Arbor. He is studying the formation of surfactant by lung tissue, the absence of which is a primary cause of death in premature infants.

These researchers have found computer methods for the storage and retrieval of their experimental results to be very helpful They are currently using standard database and statistical packages on the University of Michigan's computer system. During one year they perform experiments using approximately 2,000 explants and record values of numerous dependent and independent variables for each one. These data must be analyzed promptly so that new experiments can be planned, based on the results of previous ones. The further power that artificial intelligence techniques could provide would enable them to perform their work more effectively. Our system should relieve investigators from the difficult tasks of trying to reduce multiple variables into a single coded value and trying to give these values consistent meaning across investigators, a recurring problem in this type of analysis. The AI system that we envision will be able to store more information, the kind that is not easily codified for use in their current statistical routines, such as graphic information. Our aim is to build a system that would be able to process their data in a more sophisticated manner, being able to perform analyses and make judgments that hitherto had to be done by the researchers themselves. This should allow the researchers to spend more time actually performing experiments and thinking about basic, theoretical questions. Like other current computer programs, the AI system will be able to perform functions that were previously impossible or impractical due to the vast amount of data that need to be analyzed or to the fine discriminations that need to be made which are beyond human perception. By using natural language input and output, the AI system should be faster and more comfortable to use.

The histologists' role in this project has been to acquaint us with their existing procedures, the types of knowledge they use, and the decisions they make in order to perform organ culture research. We have been attending their weekly progress meetings and interviewing them individually. In addition to developing our AI system, we have given them consultation with regard to the use of existing computer programs on the University of Michigan computer system and the purchase of computer hardware for their department.

C. Progress Summary

Our main emphasis so far has been on the natural language processing aspects of this project. We have written programs using Interlisp at SUMEX that read typed, natural language input and perform a morphological analysis of the words. Due to the large number of Latin words that are used in anatomy, we have defined both English and Latin morphological decomposition rules. For syntactic and semantic analysis we have obtained and tested a system of Interlisp programs from Bolt, Beranek and Newman which uses a compiled semantic augmented transition network grammar. We have created a dictionary of anatomical terms and are continually improving it. Currently, we are adding to it the features used by the BBN grammar. We still need to deepen the analysis of language embodied in our programs. We would like to move beyond the morphological and syntactic levels and attempt to capture the conceptual and intentional aspects of the language of histologists. We plan to investigate programs and approaches that other members of the SUMEX community have developed or use.

We have begun work on the computerized representation of other knowledge used by histologists, such as that regarding different culture media and tissue staining techniques. We are fortunate to have on this project Dr. Kahn, who is active in the Tissue Culture Association and a member of its Committee on Chemically Defined Media. This committee has been working for the past several years to compile and index, using computer files and programs, the components of and the bibliographic references for synthetic culture media. We plan to incorporate their compilation into the database that we are currently designing. We then hope to create a similar database for tissue stains.

We have not done a significant amount of work on the image analysis side of the project this year. We have spent most of our time waiting for new hardware to help us in this regard. The Anatomy Department has recently purchased a Quantimet 720 Image Processor which contains hardware to digitize microscope slides and perform basic image analysis functions. It also includes a PDP-11 computer and peripherals for storing images and further processing them. We plan to develop image analysis programs on this machine and at a later time interface them with our SUMEX programs. We also hope to have available soon a very fast image analysis computer that the Environmental Research Institute of Michigan is completing. We tested successfully on its prototype, programs for detecting prostate alveoli from digitized microscope slides.

E. Funding Support Status

Funding for the AI related portions of this project has been supplied solely by the University of Michigan, through the Mental Health Research Institute. An application for a grant from NIH was recently rejected. A revised application may be prepared, if we feel we can meet the objections of the NIH Study Section, which felt that the "goal is worthy and the computer areas that are proposed for implementation are of great usefulness and worth pursuing," but that the project as outlined in the proposal is overly ambitious.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations, Interactions, and Sharing

The interaction and software sharing tools provided by SUMEX, such as SNDMSG, FTP, and protection mechanisms, were very useful to us in obtaining the ATN grammar compiler system from Bolt, Beranek and Newman. We found SNDMSG to be better than the mail or even the telephone for quick communication about our desires and problems, especially since our primary contact at BBN, Dr. Richard Burton, travels frequently but is almost always near a terminal. We used FTP to transfer the files from the BBN TENEX computer site to SUMEX. Dr. Burton was able to actually test the programs on the SUMEX system to ensure their proper functioning. We were surprised at the relatively little effort it took to transfer a system of this size and to obtain a working version. In addition, the ease with which we can communicate should enable us to continue our software sharing when new versions of these programs are developed

We have also had interactions with the SUMEX staff regarding minor Interlisp and TYMNET problems. We have found these staff consultants to be very helpful.

B. Critique of Resource Management

We are very pleased with the overall reliability and operation of the SUMEX system. Of the three years that we have been SUMEX users, this year has been the best in terms of reliability and response time. Now that the problem with dropped TYMNET characters has been solved and there is available a 120 cps TYMNET line, we find the editing and formatting facilities of SOS, PUB, SNDMSG, and TYPE to aid us in our formal and informal writing and to increase our productivity of reports and correspondence. We continue to find the SUMEX staff friendly and helpful. Our only criticism is a shortage of filespace. For example, we have one file that by itself takes approximately 100% of our allotted space.

III. RESEARCH PLANS (8/78 - 7/81)

A. Long Range Project Goals and Plans

Our long range technical goals have been described in the above summary. There are obviously many things to be done and some selection will need to be made. In the absence of expanded resources we will for the present concentrate on refining the language analysis systems already programmed, since these will eventually need to be extended to encompass a larger vocabulary and richer syntax. We will also work on the problem of converting images of tissue section into machine storable form, attempting to make this process more rapid and economically viable. We will also, as resources permit, develop the databases of media and of stains, both by enlarging them and by attempting to develop more useful structures for them. We feel that these projects will have some payoff even without the existence of the entire system envisioned.

B. Continuation of SUMEX Use

We have been using a rather small portion of the SUMEX resource, but have found it to be extremely valuable in our work; in fact, it has been absolutely essential. We wish to continue at approximately the same level until additional funding is obtained for the project. Our only serious limitation is in our allotment of disk space. Now that we have developed some software and compiled the beginnings of a dictionary our file space is quite cramped, and we would like to request a modest increase in our allocation.

Our programs are written in Interlisp, and SUMEX is the only facility to which we have access which supports this language. Perhaps the most important aspect of our association, however, is the opportunity SUMEX provides us for contact and interaction with other members of the AI community around the country.

We feel that we have made progress on an interesting problem which is highly relevant to the SUMEX-AIM mission. The descriptions of this work have been given in earlier sections of this report. Our hope is that this work will justify our continued access to the SUMEX facility.

C. Needs and Plans for other Computational Resources.

To develop our tissue culture model in its entirety, we need hardware for image processing. We believe that the Quantimet 720 computer recently purchased by the department of our medical collaborators can fulfill our needs in this regard. One use of the Quantimet which we foresee, is as a satellite computer to SUMEX, providing our SUMEX programs with pre-processed image data.

D. Recommendations for Future Community and Resource Development

Our basic recommendation to the SUMEX staff is to continue to provide high quality service as it has been doing. However, to guide management in selecting enhancements may we suggest the following.

1) The number of different models of display screens which the TV program can handle should be expanded so that more users can take advantage of screen editing.

2) To facilitate further software sharing the bulletin board system should be modified so that SUMEX users can describe software they have developed which might be of use to other projects. These descriptions should include for each program a person who can be contacted for more information. The current SUMEX system does a good job in sharing utility programs and documenting complete systems (such as MYCIN), but not programs that are more application oriented. While Stanford users are often able to obtain needed information by word of mouth, network users are not.

3) Any means for enhancing interactions among network users would be a great boon. For example, LINKing may not be terribly useful to Stanford users, but it could be more valuable to network users if it were more a matter of routine.

4) Along the lines of creating a greater feeling of community, perhaps an AI news service would be useful. This would inform users of current news of interest to the AI community as a whole, not just SUMEX-AIM news. Announcements of publications of important books or reviews, chess matches, government funding decisions, and so forth would be candidate items. Basically, news of the sort found in the SIGART Newsletter would be appropriate; this service would fill the gaps between Newsletters. It would not be necessary to aim at exhaustive coverage; any contributions at all would be useful.

5) Now that the bulletin board system has be in use for some time, perhaps it would be appropriate to review its design. We find it too cumbersome, and so tend to use it less than we might.

4.4 PILOT STANFORD PROJECTS

The following are descriptions of the informal pilot projects currently using the Stanford portion of the SUMEX-AIM resource pending funding, and full review and authorization.

4.4.1 GENETICS APPLICATIONS PROJECT

Computer Science Applications in Genetics

Prof. L. L. Cavalli-Sforza
Department of Genetics
Stanford University School of Medicine

I. SUMMARY OF RESEARCH PROGRAM

I.A. TECHNICAL GOALS

We are interested in understanding the role of diseases in shaping the geographic distribution of human genes. We observe a great deal of geographic variation but are still almost completely in the dark about its causes. Some clear examples have been given in the past that have linked the geographic distribution of a gene with that of a specific disease: sickle cell anemia is the best known example. In other cases the correlation was with a specific custom (e.g. lactose tolerance and the dietary use of untransformed milk).

We have been interested at the beginning in building gene frequency maps and will report here on the state of the art and the first applications of the technique.

I.B. MEDICAL RELEVANCE

Certain genetic diseases show conspicuous geographic and ethnic variation (examples: cystic fibrosis and phenylketonuria are highest in Caucasians; Tay-Sachs is found almost only among Ashkenazi Jews; gluten intolerance reaches frequencies of 1/300 in W. Ireland, and so on) - the causes are usually unknown. It could be that chance (random genetic drift) is involved.

When specific customs, infectious diseases, or pollutants can be pinned down as causal antecedents, there is usually room for preventive action. Specific interest is today given to the possibility that certain diseases may be determined by the interaction of specific stimuli on specific genotypes (e.g. emphysema in certain antitrypsin alleles; heart condition and arteriosclerosis in certain disturbances of lipid metabolism). Many of such conditions may be detectable by the study of geographic distributions and the correlations of genes and diseases.

I.C. PROGRESS SUMMARY

By far the major project this year has been the construction of GENE FREQUENCY MAPS. Alberto Piazza and Paolo Menozzi have collaborated on this project. There are some programs available which compute isopleths, but the underlying principles are either insufficiently specified, or do not correspond well enough with our needs. The data from which curves are to be built, gene frequencies, are affected by sampling error. Thus a process of surface fitting is necessary; a simple interpolation, e.g. by splines is not adequate. Moreover, attempts in different directions in the last year have convinced us that one cannot use the same fitting method for the whole surfaces but the BEST LOCAL STRATEGY OF FITTING depends on the pattern of data actually available in the neighborhood of each point to be interpolated. This is critical especially for marginal regions. We have therefore developed a program in which a set of rules is given for the strategy to be chosen, and one out of a few different strategies is selected for each point to be interpolated in a network of an appropriate mesh. The mesh is chosen on the basis of computer time vs. desired resolution. Ordinarily, more points are necessary for drawing good isopleths than these fitted points that from the nodes of a relatively rough network. We have found in earlier work that drawing isopleths followed by shading the areas defined by isopleths is possible, but inefficient. Therefore, we have at the moment avoided computing the isopleths, and use instead splines under tension to generate a mesh of the desired high degree of resolution and which goes through all the previously calculated modes of the rougher network. The fine network generated by splines can be used for instance for output on TV screen, and this method has so far been giving us the most satisfactory results.

The maps we generate are, we believe, more satisfactory, and certainly more objective and far less time consuming than hand constructed maps. This, however, is not the major scientific benefit to be obtained from map construction by computer. We think the following is, instead the major result of interest. For the study of some factors of evolution, in particular migration, it is essential to appropriately average data over as many genes as possible. We take as weighted averages of gene frequencies the linear functions called "principal components" (corresponding to leading eigenvectors of the dispersion matrix). These have the advantages of maximizing the variation between populations. Plotting isopleths of the leading principal component, instead of gene frequencies, one can summarize the information contained in all gene frequencies known. In this way we have been able to show very recently that migrations as old as that occurring 9000 to 5000 years ago, with the diffusion of farmers from the Near East are still visible - and in fact are the dominant component - in the gene frequency map of Europe. Similar conclusions can be obtained from gene frequency maps of single genes but they are far less clear and convincing, since only a few genes show the expected gradients in a clear fashion, and only by appropriately summing the evidence from many genes does a clear picture emerge.

For this purpose of computing principal components however, one needs to have compact matrices of gene frequencies x populations with no missing items. Most real data are very far from this ideal condition. Having developed a satisfactory program to make gene frequency maps, we could then proceed as follows: make for each gene a map; sample for all genes the map values at a specified network points; compute principal components from this sample of interpolated values. The procedure was validated in a case in which we could obtain principal components independently, (on a real data matrix without missing items). We were thus able to generate maps of principal components of gene frequencies in spite of the extreme incompleteness of the raw data.

The results we have obtained by this new technique have gone beyond our expectation. We find that the amount of information we can synthesize by principal components tends around values of 30% for the first, 20% for the second and 10% for the third. Thus we can synthesize 60% of the information by plotting the first three principal components in color (on a color TV screen produced by Grinnell of Palo Alto). This gives a very clear picture of the similarities between populations, as they can be gauged by synthesizing the maximum amount of information which can be stored in 3 dimensions (see figure 1).

II. INTERACTION WITH THE SUMEX AIM RESOURCE

A) We have asked the collaboration of all laboratories involved in HLA research so as to get data - published and unpublished - for a gene frequency bank. These data, as received, are stored in files. We plan to publish summaries at intervals (every 1-2 years) for distribution to collaborating laboratories, and publish them. The data will also be important for keeping up-to-date maps of HLA.

B), C) Nothing important.

III. RESEARCH PLANS

IV.A. LONG TERM PROJECT GOALS AND PLANS

Recently a big volume collecting data on gene frequencies for the most important human genes (excluding HLA and Gm) has been published by Mourant and others. This supersedes earlier attempts I had started to collate the same material, which had generated a less complete collection. Mourant's work has been that of a lifetime, but there were fears he might not be able to finish and publish his work. The second edition of his book which finally appeared in 1976 has been overdue for at least 6 or 7 years. The data thus available, from over 4000 original papers, deserve a full analysis. We have, I believe, made a good start to test them for one evolutionary factor, migration. This is a preliminary step of interest per se, in that it allows to reconstruct, for instance, population history; it is also an important basis for the study of other evolutionary factors. Having first eliminated the effect of migrations one can look at residual effects due to natural selection (and therefore diseases) and drift. Having built satisfactory maps of gene frequencies one can find selective responses to climatic conditions by correlating gene frequencies and climatic data. We have had remarkable results using anthropometrics instead of gene frequencies in a related project; an initial trial with gene frequencies has been moderately successful. We can also look at the correlation of the geographic distribution of diseases, as available from WHO statistical summaries and other sources, with that of single genes. There are also atlases of cultural data (e.g., Murdock's Ethnographic atlases) and an archive of languages which is computerized (c/o Stanford Department of Linguistics) which can be used to supply maps of cultural influences and which may be directly related to gene exchange, given that cultural and genetic migration are some time simultaneous, and that culture is another aspect of the environment in which we live. These are just indications of the lines of research we want to follow, on a long term basis for understanding more about factors of genetic differentiation and of environmental factors (including certain cultural ones) and their impact on human diversities with special consideration to health factors. Naturally, we are aware that the finding of a correlation is only a first step in the detection of causal antecedence, and there are many ways in which correlations can fool us if accepted initially or without further work at other levels. The study of correlations is to be thought, however, as a screening method for detecting areas when further research may be useful. In the way we practice it, moreover, we have several ways of showing that a correlation is spurious. If we look at the correlation of a gene frequency with say the incidence of an infectious disease, (or a climatic variable), we expect the correlation to be found not only at a global level, but even more clearly in different world subareas. Moreover, our work on migration and other evolutionary factors allows us to eliminate other causes of variation of gene frequencies, namely all those that are common to many genes and not idiosyncratic to some. The latter is instead likely to be true of most instances of natural selection determined by disease.

On a shorter-term basis, we are interested in first improving our present program of map building. They already represent a first attempt at automation of a complex series of decisions in map building. We can probably improve greatly by recourse to Dendral. If we are successful in making the building of maps more precise and also more efficient we can hope to enlarge the range of problems which can be considered. From a technical point of view we can improve maps by simultaneously fitting gene frequencies for all alleles at a locus (whenever there are more than two alleles) because we can by simultaneous fitting take account of the restriction that all allele frequencies must sum to one. At the moment we do not consider this constraint. A more important program is that of studying drift by simultaneous fitting of several gene frequency maps. Populations which are of smaller size and are therefore expected to have more drift will show greater local variation. This is expected to show as local anomalies of the gene frequency map FOR SEVERAL GENES. A simultaneous fitting is therefore the simplest way of looking for such anomalies.

It is however more satisfactory to entrust the computer with as many as possible of the decisions to be taken, and the interaction would if anything be used as a temporary step for understanding more about the problems which arise and whose existence we may be at the moment unable even to anticipate. We are now trying to secure the help of a computer scientist to work on a more permanent basis on this problem. We thus hope to be able to make use of more sophisticated techniques, e.g. Dendral, to solve our problems.

We have been using in the past Sumex also for other projects; one of these, perhaps the biggest is Dr. Ammerman's storage-analysis-retrieval system of neolithic Calabrian data. But at present the load we have as compared with the number of pages available is such that we can only do one type of thing at a time, storing everything else on tape in the meanwhile. I had to ask Dr. Ammerman to find room elsewhere, in Sumex or another computer, even though I am vitally interested in his results and am planning to collaborate with him on their analysis when the storage will be finished. Other projects going on intermittently involve techniques of multivariate analysis. We are especially interested in developing new techniques for answering specific problems, for instance 1) the simultaneous fitting of means, in addition to variances and covariances. At the moment, only the latter are considered. One result we have had is that parameters derived from variances/covariances only do NOT fit means (e.g., the means of behavioral characteristics, like IQ, for adopted children given those of parents). This may be due to an insufficiency of the model, or of the present fitting process, or of both; 2) the search for new synthetic ways of expressing multivariate data which maximize specific quantities: e.g. a function maximizing the correlation between adopted child and biological parent (which will express genetic factors) and one maximizing the correlation between adopted child and adoptive parent (which will express cultural factors).

I have been using multivariate analysis extensively and am convinced that, if properly used, it is the major help that artificial intelligence can give when complexity is due to the multiplicity of traits being considered. There are probably few other people interested in multivariate analysis in the Sumex community, since all that is available is a Factor Analysis program in SPSS and there are no programs for handling matrices (inversion and especially spectral analysis). Although there has been considerable increase of interest (especially among psychologists) multivariate analysis techniques remain relatively unused. I find that several of my problems had to await a really satisfactory solution for a long time until I tried classical techniques of multivariate analysis, the "principal components" we used for building synthetic gene frequency maps. Basically, multivariate analysis simplified problems by reducing the dimensionality of the data. If it were limited to this, one might think of it as a "mechanical" kind of artificial intelligence. There is, however, a dynamic aspect to be developed, in that there are many options and the choice of one could be made a more intelligent and automatic process than it is now.

There are many things we would have been unable to do without Sumex, but we have also been under constant shortage of space. I doubt we will be able to extend our work on maps if we are unable to obtain more space.

III.C. NEEDS AND PLANS FOR OTHER COMPUTATIONAL RESOURCES

We have a couple of projects at the planning stage and for which we many need more access to a computer. As yet I am not aware whether they will be of A.I. stype or involve more conventional computer usage. Accordingly, in the financing of the project (now been submitted to NSF) we ask for a small size minicomputer. Another resource of interest is a color TV screen, of the type we have used by kind permission of Prof. R. Lyon of the Earth Sciences Department. We are considering the possibilities of acquiring one, if we can find the financing.

4.4.2 QUANTUM CHEMICAL INVESTIGATIONS

Theoretical Investigations of Heme Proteins, Opiate Narcotics, Chemical Carcinogens, Drug Metabolism, DNA Structure and Intercalating Drugs

Dr. Gilda Loew
Molecular Theory Laboratory
Department of Genetics
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

The major theme of research in the Molecular Theory Group is the application of the techniques of theoretical chemistry to a variety of biomedical problems. To this end a number of large scale computer programs which embody these techniques are utilized to characterize the electronic structure, conformation, interactions with appropriate receptors, and electromagnetic properties of biologically important molecules. The SUMEX-AIM resource is used to characterize starting geometries of appropriate molecular systems used as input to the large scale programs and to calculate electromagnetic properties from the output of such programs. It is also used for rapid energy calculations of opiate peptide conformations. This judicious use of the facility thus enhances the productivity and efficiency of all of of our research projects. A brief description and status of specific areas of research are given below.

A. Structure Activity Studies of Opiate Narcotics

Many classes of opiates have been studied including the newly discovered endogenous peptide opiates. Similarities and differences among these classes have been delineated and molecular properties related to agonist/antagonist potency ratios identified and characterized. (See references 1-3 for the latest of 11 publications in this area). This work is in its fourth year of support from the National Institute of Drug Abuse. Collaboration with medicinal chemists and clinical pharmacologists interested in drug design is under way.

B. Drug Metabolism and Activation of Chemical Carcinogens by Cytochrome P450

Models of the active site of this ubiquitous metabolizing enzyme in its biologically active state have been made and electromagnetic properties predicted. In addition, structure activity studies of classes of substrates, specifically general anesthetics and polycyclic aromatic carcinogens, have been made which relate molecular reactivity to substrate efficacy and metabolite distribution (see references 4-9). This work has been modestly supported in the past by NSF and ACS and is currently being supported by a new grant from NIH (GM) and a one year contract from NCI.

C. Chemical Carcinogens

The goal of this study is to help identify and calculate properties of parent compounds related to their carcinogenic potency and hopefully to ultimately predict which members of given classes of chemical carcinogens will be active. Models for metabolic activation, mode of action and interaction with DNA are part of this project which is funded by a one year contract from NCI (same as above).

D. Structure Function and Electromagnetic Properties of Heme Proteins and Related Metal Organic Complexes

Using a reasonable model as input, large scale molecular orbital programs are used to calculate the electronic structure and conformation of the active site of heme proteins and the mode of binding of a number of biologically relevant ligands such as CO, O2 and CN-. Using the calculated electronic structures and conformations, a set of auxiliary programs are used to calculate measurable electromagnetic properties such as quadrupole splitting observed in Mossbauer resonance spectra, g values and hyperfine splittings in electron spin resonance spectra, and magnetic moments. This close connection between observables and basic molecular structure enhances the usefulness of experimental data in inferring such fundamental molecular properties as the nature of metal-ligand binding and how small changes in conformation at the active site affect biological function (references 10-12). This work has had modest continual support from NSF for the past twelve years

E. DNA Structure and Interaction with Intercalating Molecules

This work is divided into two parts: Investigations of the various modes that DNA can "kink" consistent with known chromatin structure and the origin of the specificity in the binding of ethidium-type intercalators into DNA components (references 13-15). This work is not currently funded.

II. INTERACTION WITH SUMEX-AIM RESOURCE

The SUMEX-AIM resource is used by the Molecular Theory Laboratory on a limited basis which allows optimum efficiency on almost all of our projects. Specifically we have used it for:

A. Interactive Data Preparation and Input for all Projects

The procedure for the characterization of heme proteins initially requires a determination of a model for the active site. The model has to be large enough to realistically describe interactions that take place at the active site but not so large that it becomes economically infeasible to perform a calculation on the molecule. Once a general model is chosen, the specific geometry for the molecule must be determined. With the aid of experimental crystal structures of related compounds, we use the SUMEX-AIM facility to interactively determine the coordinates of the best initial approximation to the geometry of the active site model. Specifically for heme proteins this involves determining the hole size of the porphin ring, the degree of nonplanarity of the porphin ring, the position of the metal atom at the center of the active site, and the nature and geometry of the axial ligands bound to the metal.

For the study of the relative properties of a family of carcinogens as a function of substituent on a given parent compound, the SUMEX-AIM facility is used to interactively determine and plot the most reasonable geometries of the various substituents depending on the parent compound. Similarly, the geometries of DNA base-pairs and organometallic transition metal complexes are calculated using SUMEX-AIM.

B. Calculation of Electromagnetic Properties of Iron Containing Proteins and Related Organometallic Compounds

We are currently performing systematic studies of heme proteins including the metabolizing enzyme cytochrome P450. The electromagnetic properties of these proteins and of synthesized model compounds which mimic the observed behavior of the proteins have been well studied experimentally in a number of instances. SUMEX-AIM is used for the calculation of these one-electron properties.

The properties that are calculated include the electric field gradient at the iron nucleus, quadrupole splitting, isotropic and anisotropic hyperfine interaction, spin-orbit coupling and zero field splitting, g values and temperature dependent effective magnetic moments. The calculated values are compared directly to experimental results obtained from published Mossbauer resonance and electron spin resonance spectra. Such a comparison determines not only the reliability with which these properties can be calculated but also gives an indication of the ability of the model of the iron active site to mimic the actual environment found in a particular compound or iron containing protein.

The major input to these properties programs is a description of the electron distribution of the compound under consideration. This description is obtained using a semiempirical molecular orbital method employing the iterative extended Huckel procedure. Such a calculation requires up to 660K core and is performed elsewhere. When the calculated electron distribution yields a set of calculated properties in agreement with observation, we have increased faith in the description of the model of the active site and can carry the model one step further to make qualitative inferences about certain properties relevant to the biological function of the compound.

C. Conformational Study of Pentapeptides

In a completely different context, we have used SUMEX-AIM to calculate the conformation of pentapeptides (enkephalins) which have been recently found to be endogenous opiates. The aim of this study was to determine in what way, if any, they can mimic the structure of prototype opiates such as morphine and meperidine. For this work, we have used a protein conformation program with empirical interaction potentials. Quantum mechanical conformation calculations of the same peptides have been performed by us elsewhere and the results of the two methods have been compared (reference 2).

III. RESEARCH PLANS (8/78-7/81)

We plan to continue research in the same general areas as our current projects: Systematic studies of iron-containing proteins; structure activity studies of opiates and other drugs; drug metabolism and activation of chemical carcinogens; structure activity studies of classes of chemical carcinogens; structure of DNA and binding of DNA components to small drugs and carcinogens; and studies of peptide conformation.

Specific judicial use of SUMEX-AIM is planned in connection with most of these projects in the same spirit as we have described for our on-going projects:

1. We plan to use SUMEX-AIM to continue to characterize reasonable molecular geometries to use as input to our large scale computer programs which calculate molecular properties and also intermolecular interactions (drug/receptor, enzyme/substrate). The flexible operating mode of SUMEX-AIM is important to the efficient and accurate preparation of input before costly time consuming production runs are submitted at another facility.

2. We plan to use SUMEX-AIM to continue our systematic study of heme proteins, specifically in the calculation of electromagnetic properties. These calculations will help to characterize the four stable states of the metabolizing heme enzyme cytochrome P450 and to study the relationship between conformation and function of other classes of heme proteins (electron and oxygen transfer proteins).

3. We plan to use SUMEX-AIM for a new study of peptide conformations using the experience gained from the investigation of peptide opiates. Specifically, in collaboration with G. Fassman we plan to investigate specific correlations between primary sequence and conformation of proteins and to do limited structure activity studies of some peptide hormones.

4. We plan to use SUMEX-AIM to help decide plausible geometries for complexes formed by the active form of chemical carcinogens with DNA components.

SUMEX-AIM has become an essential component of our research procedure. The continued use of SUMEX-AIM is central to the efficient implementation of our research plans. We are grateful for past support and feel justified in our request for continued support.

REFERENCES

1. Loew, G. H., Berkowitz, D. S., Quantum Chemical Studies of N-substituent Variation in the Oxymorphone Series of Opiate Narcotics, J. Med. Chem., 21, 101 (1978).

2. Loew, G. H. and Burt, S. K., An Energy Conformation Study of the Resemblance of Met-Enkephalin and D-Ala2 Met-Enkephalin to Rigid Opiates Using Empirical and Quantum Mechanical Methods. Proc. Natl. Acad. Sci. 75, 17, (1978).

3. DeGraw, J. I., Lawson, J. A., Crase, J. L., Johnson, H. L., Ellis, M., Uyeno, E. T., Loew, G. H., and Berkowitz, D. S., Analgesics. I. Synthesis and Analgesic Properties of N-Sec-Alkyl and N-Tert-Alkylnormorphines, J. Med. Chem. (1978), in press.

4. Hjelmeland, L., Loew, G. H., The Electronic Structure of Peracids: Functional Models for Cytochrome P450, Tetrahedron, 33, 1029 (1977).

5. Loew, G. H., Kert, C., Hjelmeland, L., Kirchner, R. F., Active Site Models for Horseradish Peroxidase Complex I and a Cytochrome P450 Analogue, J. Amer. Chem. Soc., 99 (10), 3534 (1977).

6. Hjelmeland, L., Loew, G. H., The Electronic Structure for Chromyl Chloride: A Functional Model for Cytochrome P450, J. Am. Chem. Soc., 99 (10), 3514 (1977).

7. Loew, G. H., Hjelmeland, L., Kirchner, R. F., Models for the Enzymatically Active State of Cytochrome P450, Int. J. of Quant. Chem., QBS, 4, 225 (1977).

8. Loew, G. H., Wong, J., Phillips, J., Hjelmeland, L., Pack, G., Quantum Chemical Studies of the Metabolism of Seven Benzo(a)pyrene, Can. Biochem. Biophys., in press (1978).

9. Loew, G. H., Phillips, J., Wong, J., Hjelmeland, L., Pack, G., Quantum Chemical Studies of the Metabolism of Seven Polycyclic Aromatic Hydrocarbons, Can. Biochem. Biopys., in press (1978).

10. Kirchner, R. F. and Loew, G. H., Semiempirical Calculations of Model Oxyheme: Variation of Calculated Electromagnetic Properties with Electronic Configuration and Oxygen Geometry, J. Am. Chem. Soc., 99, 4639 (1977).

11. Loew, G. H. and Kirchner, R. F., Semiempirical Calculations of Model Deoxyheme. Variation of Calculated Electromagnetic Properties with Electronic Configuration and Distance of Iron form the Plane, Biophys. J., 22, 000 (1978).

12. Loew, G. H. and Kirchner, R. F., Binding of O2, NO, and CO to Model Active Sites in Ferrous Heme Proteins: Ligand Geometry, Electronic Structure and Quadrupole Splittings, Int. J. Quan. Chem. QBS, in press (1978).

13. Pack, G. R., Muskavitch, M. A., Loew, G. H., Kinked DNA: Energetics and Conditions Favoring its Formation, Biochim. Biophys. Acta, 478, 9 (1977).

14. Pack, G. R.and Loew, G. H., The Origins of Sequence Specificity of Ethidium Nucleic Acid Intercalation, Int. J. Quant. Chem. QBS, 4, 87 (1977).

15. Pack, G. R. and Loew, G. H., Origins of the Specificity in the Interaction of Ethidium into Nucleic Acids: A Theoretical Analysis, Biochim. Biophys. Acta, in press (1978).

Appendix I

AIHANDBOOK OUTLINE

E. A. Feigenbaum
A. Barr
Computer Science Department
Stanford University

This is a tentative outline of the Handbook. Articles in the first eight Chapters are expected to appear in the first volume. A list of the articles in each Chapter is appended.

I. Introduction
II. Heuristic Search
III. AI Languages
IV. Representation of Knowledge

V. Natural Language Understanding
VI. Speech Understanding
VII. Applications-oriented AI Research
VIII. Automatic Programming

IX. Theorem Proving
X. Vision
XI. Robotics

XII. Information Processing Psychology
XIII. Learning and Inductive Inference
XIV. Problem Solving, Reasoning, and Planning

I. INTRODUCTION
- A. The AI Handbook (intent, audience, style, use, outline)
- B. Overview of AI
- C. History of AI
- D. Philosophy of AI
- E. AI and Society
- F. An Introduction to the Literature in the field

II. Heuristic Search
- A. Overview
- B. Problem representation
  1. Overview
  2. State space representation
  3. Problem reduction representation
  4. Game trees
- C. Search
  1. Blind state space search

- 2. Blind search of an and/or tree
- 3. Heuristic search in problem-solving
    - a. Basic concepts in Heuristic Search
    - b. A*: optimal search for an optimal solution
    - c. Relaxing the optimality requirement
    - d. Bidirectional search
    - e. Heuristic search of an and/or graph
    - f. Hill climbing
- 4. Game tree search
    - a. Minimax
    - b. Alpha-beta

D. Example Programs
- 1. Logic Theorist
- 2. GPS
- 3. Gelernter - geometry
- 4. Slagle & Moses - Integration
- 5. STRIPS
- 6. ABSTRIPS

III. AI Languages

A. Overview of AI Languages (Historical)

B. Comparison of AI Languages

C. Early list-processing languages (SLIP, IPL, SNOBOL)

D. Current languages/systems
- 1. LISP, the basic idea, INTERLISP
- 2. SAIL/LEAP
- 3. POP-2, POP-10
- 4. QLISP (mention QA4)
- 5. PLANNER
- 6. CONNIVER
- 7. Object-oriented languages (ACTORS, SMALLTALK, SIMULA)
- 8. FUZZY (LeFaivre@Rutgers)
- 9. QA3/PROLOGUE
- 10. PC programming languages

IV. Representation of Knowledge

A. Survey of representation techniques

B. Issues and problems in representation theory

C. Representation Schemes
- 1. Predicate calculus
- 2. Semantic nets (Quillian, LNR, Hendrix)
- 3. Production systems
- 4. Procedural representations (SHRDLU, actors, demons)
- 5. Semantic primitives, Componential analysis (Fillmore, Schank, Wilks)
- 6. Direct (Analogical) representations
- 7. Higher Level Knowledge Structures
    - a. Frames, Scripts, The basic idea (Bartlett, Minsky, Abelson)
    - b. KRL-0, MERLIN
    - c. Others: FRL, OWL, Toronto

D. Discussion and Comparison of Representation Schemes

V. Natural Language
- A. Overview - History & Issues
- B. Grammars
  1. Review of formal grammars
  2. Transformational grammars
  3. Systemic grammars
  4. Case Grammars
- C. Parsing techniques
  1. Overview of parsing techniques
  2. Augmented transition nets, Woods
  3. CHARTS - GSP
- D. Text Generating systems
- E. Machine Translation
  1. Overview & history
  2. Wilks' machine translation work
- F. Natural Language Processing Systems
  1. Early NL systems (SAD-SAM through ELIZA)
  2. PARRY
  3. MARGIE
  4. LUNAR
  5. SHRDLU, Winograd

VI. SPEECH UNDERSTANDING SYSTEMS
- A. Overview (Includes a mention of ac. proc., blackboard)
- B. Design Considerations for Speech Systems
- C. The Early ARPA speech systems
  1. DRAGON
  2. HEARSAY I
  3. SPEECHLIS
- D. Recent Speech Systems
  1. HARPY
  2. HEARSAY II
  3. HWIM
  4. SRI-SDC System

VII. Applications-oriented AI research
- A. Overview of AOAIR
- B. CHEMISTRY
  1. Mass Spectrometry (DENDRAL, CONGEN)
  2. Organic Synthesis (Wipke@UCSC)
- C. MEDICINE
  1. Overview
  2. MYCIN
  3. Glaucoma
     - a. CASNET,
     - b. IRIS
  4. DIALOG, INTERNIST II
  5. PRESENT ILLNESS (MIT)
  6. DIGITALIS (MIT)
  7. TEIRESIAS
- D. MATHEMATICS
  1. REDUCE

        2. MACSYMA
        3. AM
    E. EDUCATION
        1. Overview
        2. SCHOLAR
        3. SOPHIE
        4. WEST, BUGGY
        5. COACH
    F. MSC.
        1. SRI Comp. Based Consultant (PROSPECTOR)
        2. RAND-RITA
        3. AI applications to Information Retrieval
        4. SU/X
        5. Management applications

VIII. Automatic Programming
    A. Automatic Programming Overview
    B. Program Specification
    C. High-level Program Model Construction
    C. Program Synthesis
        1. Overview
        2. Techniques (Traces, Examples. Natural Language, TP)
            a. Traces
            b. Examples
            c. Natural Language
            d. Theorem Proving
    D. Program optimization techniques
    E. Programmer's aids
    F. Program verification (see Article IXD5)
    G. Integrated AP Systems

IX. THEOREM PROVING
    A. Overview
    B. Predicate Calculus
    C. Resolution Theorem Proving
        1. Basic resolution method
        2. Syntactic ordering strategies
        3. Semantic & syntactic refinement
    D. Non-resolution theorem proving
        0. Overview
        1. Natural deduction
        2. Boyer-Moore
        3. LCF
    E. Uses of theorem proving
        1. Use in question answering
       [2. Use in problem solving]
        3. Theorem Proving languages
        4. Man-machine theorem proving
        5. In Automatic Programming
    F. Proof checkers

X. VISION
- A. Overview
- B. Image-level processing
  - 1. Overview
  - 2. Edge Detection
  - 3. Texture
  - 4. Region growing
  - 5. Overview of Pattern Recognition
- C. Spatial-level processing
  - 1. Overview
  - 2. Stereo information
  - 3. Shading
  - 4. Motion
- D. Object-level Processing
  - 1. Overview
  - 2. Generalized cones and cylinders
- E. Scene level processing
- F. Vision systems
  - 1. Polyhedral or Blocks World Vision
    - a. Overview
    - b. COPYDEMO
    - b. Guzman
    - c. Falk
    - d. Waltz
    - e. Navatya
  - 2. Robot vision systems
  - 3. Perceptrons
  - 4. etc.

XI. ROBOTICS
- A. Overview
- B. Robot Planning and Problem Solving
- C. Arms
- D. Present Day Industrial Robots
- E. Robotics Programming Languages

XII. Information Processing Psychology
- A. Overview
- B. Memory Models
  - 1. Overview
  - 2. EPAM
  - 3. Semantic Net Models
    - a. Quillian & Collins
    - b. HAM-ACT (Anderson & Bower)
    - c. LNR ASNs
  - 4. Production Systems a Memory Models (Newell, Moran, ACT)
  - 5. Higher level structures (Schemas, scripts & Frames)
- C. Human Problem Solving
- D. Behavioral Modeling
  - 1. Belief Systems
  - 2. PARRY
  - 3. Conversational Postulates (Grice, TW)

4. Politics, Abelson R. and J. Carbonell, Jr.,

XIII. Learning and Inductive Inference
- A. Overview
- B. Simple Inductive Tasks
  1. Sequence Extrapolation
  2. Grammatical Inference
- C. Pattern Recognition
  1. Character Recognition
  2. Other (e.g. Speech)
- D. Learning Rules and Strategies of Games
  1. Formal Analysis
  2. Individual Examples of Games-learning programs
- E. Single Concept Formation
- F. Multiple Concept Formation: Structuring a Domain (AM, Meta-DENDRAL)
- G. Interactive Cumulation of Knowledge (TEIRESIAS)

XIV. Problem Solving, Planning & Reasoning by Analogy
- A. Overview of Problems Solving
- B. Planning
  1. Overview (pointers to discussions in Search, Robotics, AI Languages)
  2. STRIPS (see IID5)
  3. ABSTRIPS (see IID6)
  4. NOAH
  5. HACKER
  6. INTERPLAN (Tate)
  7. Rieger's inference engine
  8. BELIEVER (Schmidt, Sridharan)
  7. QA3 (see IXE1)
- C. Reasoning by Analogy
  1. Overview
  2. Evans's ANALOGY Program
  3. ZORBA
  4. Winston (see Learning)
- D. Constraint relaxation
  1. Waltz (see Vision)
  2. REF-ARF
- E. Game playing
  (This overview must point to work in search and discuss
  GP programs of various misc. sorts)

Appendix II

DIGITAL COMMUNICATIONS AND SCIENCE

Digital Communications and the Conduct of Science
THE NEW LITERACY

Joshua Lederberg, Senior Member, IEEE
Department of Genetics
Stanford University

Abstract: This essay is a personal perspective on the emergence of a new form of communication, optimistically called the 'eugram'. This form is based on the convergence of economical digital communications with computer aided facilities for file management, and protocols to facilitate the interconnection of users separated both in time and space. The eugram is contrasted with the telephone, with the latter's demands on instant availability and the subjugation of the user to an almost uninterruptible stream of data. The eugram is expected to increase the thoughtfulness of communication, the return of literacy in the efficient and precise use of language, and to enhance scientific discourse in many other ways.

Introduction

Computer communication networks provide new tools and opportunities for the scientific community to share scarce computer-based resources. They permit a new form of informal communication between scientists and often provide motivation and reward for timely sharing of research results. In addition computer-based support to large distributed segments of a scientific community is made possible via users and computers interconnected by computer controlled networks.

Today the most significant and useful form of computer communication is based on packet switched technology which has been reduced to practice in daily support of some portions of he scientific community.

Two key elements of this technology base are:

- Computer-based user-user message capability, i.e., electronic mail plus the computer-management of text data.

- Sharing in the development, refinement and use of large, complex computer knowledge-based systems particular to a segment of science, which would not otherwise be widely available.

This essay is written from the perspective of an enthusiastic USER of packet switched communications. The system itself is here regarded as a black box that accomplishes efficient transfer of digitally encoded information in near-real time among terminals that interface both to human users and to computer-manageable files. The economical integration of user, file,processor, and distance-indifferent communication link is the novel capability of what I shall call a EUGRAM system. EUGRAPHY thus embraces not only electronic despatch of mail but also a panoply of computer-augmented text handling tools and protocols. This account is informed by my experience over the last five years in the development of the SUMEX-AIM community for research in artificial intelligence related to biomedical science. However, it will be primarily concerned with the expected impact of, and needs for, the elaboration of EUGRAPHY in the conduct of scientific research generally over the next 25 years.

Conduct of Science: Computers and Communications

The claim of science to universal validity is supportable only by virtue of a strenuous commitment to global communication. In the spatial domain, the canon of publication insists upon public awareness and criticism of avowedly new knowledge. This enforces the reliability of empirical reports and assembles them into common models of a real world. In the temporal domain, the archiving and retrieval of information sustains the discipline of novelty -- assuring that we acknowledge, so as to be able to extend, the boundaries of 'human', i.e., universal knowledge.

The past twenty years have witnessed a growing self-consciousness about the structure of scientific activity, impelled in part by Malthusian concerns over the long term implications of a geometric increase at 0.25 db/yr: a ten-fold expansion over the 40-year typical career of the scientist. Much more has been written than implemented about means of helping scientists keep up with the "information-explosion". One must acknowledge the utility of recent introductions of literature-searching and alerting services, many of which crucially depend on computer support and EUGRAM-like communications. On the other hand, it will probably be the cost-explosion of print media for scientific publications [1] that proves to be a more immediately compelling motive for fundamental reexamination of our methods of scientific documentation and communication. Designs for solving these problems -- reviewed long since [2] -- must take into account that the media for communication also play a crucial role in quality control in science. The filtering procedures of the 'refereed journal' support the selection both of worthwhile reading, and of the workers whose established performance entitles them to the privileges of academic positions and social subsidy for their research.

Perhaps on account of these latter concerns, most of my colleagues in biomedical research would be loath to adopt many changes in the present system of print publication. In practice, frequent personal encounters [3] facilitated by grant funding, jet aircraft and invisible colleges [4,5] seem to play an increasingly important role in the exchange of information within scientific specialities, but without any systematic inquiry as to the costs, efficiency, and equitability of these modes. Nevertheless, no piece of work, no claim to priority, is authentically recorded until it has appeared in public print in a respectable refereed journal. The long-distance telephone surely has its role also, but more for operational detail than serious intellectual discourse; and the use of the mails is as idiosyncratic as is the performance of the US Postal System, with the notable exception of the exchange of xerocopied preprints of forthcoming publications.

In the face of this inertia, one should be skeptical about the marketability of new systems like EUGRAPHY, regardless of their technical virtues. Indeed, scientists may be the last to adopt them on a comprehensive scale, except for demonstrations that may arise from a) computer science, b) research management, c) military requirements, d) the ever graver collapse of conventional mails, and e) business applications like EFTS. With respect to c), we of course owe a great deal to the ARPANET as showing the way, and with the potential for a spillover into civil technology perhaps comparable only to jet engine. The sheer economy of EUGRAPHY, and the diffusion of microprocessors and displays into the laboratory and into everyday life, are bound to force an encounter with the challenges of new systems despite the traditional conservatism of the scientific establishment (with respect to its own way of doing business, and its attention to change outside the academic discipline [6]). Nevertheless, the history of the medical and engineering sciences both show many instances where a reluctant marriage of theoria and praxis has engendered major enrichments of the basic sciences.

All the above notwithstanding, our own experience with EUGRAPHY at SUMEX-AIM has been extraordinarily good. Individual users of course rely upon it routinely for access to computer processing. More surprising was the utility of EUGRAPHY for research management, involving the exchange of texts even over relatively short distances -- offices down the corridor or in nearby buildings. This phenomenon has provoked introspections about EUGRAPHY as a qualitatively different method of interpersonal communication from conversation, the telephone, the handwritten memo, the dictated letter or the published report, and some speculations about the further evolution of EUGRAMs as part of scientific communication.

Comparing the EUGRAM with the Telephone [7,8]

When telephone usage is limited to a few calls per day, and the connecting parties are reliably locatable, the telephone may indeed fulfill its image of instant, spontaneous communication. In current practice, beleaguered by time zone shifts, lunch hours, conferences, and competing calls, the reality of phone usage is exemplified by the employment of secretaries to make and receive the calls. The very instantaneity of the phone connection generates a queueing problem that defeats the basic motive. In due course, the two-way conversation may disappear, to be replaced by messages stored on tape recorders. The information density of speech may be viewed as very low, or very high, depending on how much of the burden is carried by the text, how much by inflection, phrasing, and other personal qualities. It may be only with respect to communications that have high affective content that audio- can compete with digital channels, and to do this well may require better than the average channel quality than is now readily available between metropolitan centers. Even here, the enhancement of literary competence might go a long way to permit the EUGRAM to compete with the song.

The EUGRAM, furthermore, has all the advantages of digital storage and accessibility to archiving, sorting and searching mechanisms that are far easier to implement, and require far less bandwidth than do voice messages. The EUGRAM itself can be composed quickly with a text-editor on the user display, where it is readily rehearsed, corrected and re-edited before being transmitted. The same EUGRAM can be fanned out simultaneously to a large number of recipients, or it can be revised and perfected through several versions with similar broadcast, or with selective distribution.

From the receiver's perspective, he has the advantage of a literate spatially oriented medium. In contrast to the time-fluent telephone, radio or TV, he has the option of perusing his mail at his own pace, of interruption, backtracing and cross-checking the text, even of marking it for reexamination and further rumination. He retains mastery of the use of his own time, and can coordinate attention to a coherently chosen set of tasks. He is liberated from the tyranny of synchronizing his own mental processes to those of the external actor. This freedom of course reduces the impact of that actor, just in proportion to the responsible autonomy it returns to the reader.

In framing responses, entire messages or selected extracts together with added comments can be forwarded to others, or returned to the sender -- lending focus to a 'discussion' and providing unambiguous texts for the development of a consensus. EUGRAMs can be filed and retrieved efficiently, or transcribed into hard copy as required. Text editors may be embellished with elaborate formatting aids, spelling correctors, even an online thesaurus to aid in composition. When quantitative calculations are in question, numbers can be mechanically copied directly from program outputs, avoiding pestiferous typographical errors. The same computer is likely to be the user's research tool and give access to shared data-bases: the EUGRAMs can then refer to common files by names that are themselves machinable. The user will also have access to other conveniences, such as desk-calculator-like programs for the checking of figures. He can even track the growing size of a EUGRAM-script (like this one) to be sure it fits into the assigned space. These word-processing capabilities can of course be consummated with hard copy sent through the mails, but with some additional effort, and the degradation of the machinability of the product at the other end.

The paradoxes of instant telephony are most manifest when several parties are involved. In our experience, several weeks prior notice (or other rigid prearrangement) has been needed to schedule teleconferences if four or more people were required simultaneously. EUGRAMs to groups are sent in real time supported by conveniences like group labels. Stored in the receiver's file areas, EUGRAMs are exchanged among an active community like SUMEX-AIM within a few hours, often within minutes. Users also remain in ready communication with each other, via their respective EUGRAM files, even when either or both have travelled away from their customary homes. Lightweight, portable terminals give any user full access to the system from any point which connects to the global telephone and other communications networks. Some facilities offer a fair amount of directory assistance, in locating and identifying the EUGRAM addresses of users; files may also be used to contain blocks of addresses that can be addressed by group names. At SUMEX-AIM, publically accessible bulletin boards are also available for broadcasting information or posting queries, without obtrusion, to a large audience. No doubt, 'junk mail' will become a problem in this medium, as it may in any other. However, the recipient has as powerful a technology for filtering unwanted messages as the broadcaster has for disseminating them. The struggle is more evenly matched, and there is then less economic incentive for abuses than applies, for example, to the distraction of one's attention by automated telephone sales technology.

Both for the management of the administrative affairs of the system, and for many of the research communications, EUGRAMs have become the preferred method of communication, provided they can be punctuated with occasional formal presentations, and more intimate encounters to help sustain the affiliations of the group. There is still plenty of personal style in the communications, and there is little problem evoking images of the warm bodies at the terminals. This intimacy can and should be supported by encouraging the occasional use of the EUGRAM system for arranging personal rendezvous. The trivial costs of such diversions are more than compensated by the enhanced efficiency of a worker who becomes adept at the use of EUGRAMs as if they were an extension of his own voice or handwriting.

EUGRAMs and Complex Communications [9,10,11]

One of the most controversial questions in social anthropology asks: "Is there a basic difference in modes of thought as between ... 'pre-scientific' and 'science-oriented', 'literate' and 'non-literate' ..." societies [12]. The controversy is complicated by the empirical difficulties of measuring the cognitive styles of individuals independent of their social interactions and of the very media whose effects are in question. The evolutionist would have to interject that a certain neurological development was a precondition for literacy and presumably would have been subject to natural selection at least during the brief interval of human history since the invention of writing. Conversely, the oral tradition made its own demands on other centers in the brain. The only question is whether these cultural patterns have been sufficiently stable and durable to have had a significant effect on the differential evolution of the human brain in different cultures.

Without going so far into the language/thought relationship, we can be categorical about the essentiality of writing for complex cognitive performances. The list -- whether an inventory of baskets or grain, or a city telephone directory -- is an externalization of cognitive activity that invites and sustains public use and scrutiny, and a form that has no effective analogy in the oral tradition. Indeed, it may have been the initial technological breakthrough in record-making preceding other forms of literature. A glance through the pages of this journal is evidence enough of the impossibility of assembling complex scientific arguments without the use of the written record. The manipulation of recorded symbols is a pale shadow of an internal cognitive imagination we hardly understand, but our most intricate intellectual exercises rely heavily on those external marks.

In many instances, it might still be possible to read a journal article over the telephone and garner some degree of comprehension of the argument even without visible records: but consider how often we have to ask simple names to be spelled out and numbers repeated in phone discourse. Imagine then communicating a computer program of more than ten instructions over the telephone! Indeed, it is precisely for the sharing of such program source texts that EUGRAMs have been most manifestly indispensable for groups like the ARPANET and SUMEX-AIM communities.

These program texts, which may reach hundreds of thousands of instructions are among the most complex records of human logical effort -- and more than any other production, the information is manifestly all in the text. However, they also typify the information content of other scientific efforts like mathematical proofs, structural analysis in chemistry, and other arguments. Some of these also resemble program sources in becoming almost impossible to criticize as written records alone, viz., without exercising them on the computer or in the laboratory. The recent demonstration of the four-color-map theorem comes to mind [13].

One of the facilities offered under SUMEX-AIM is the CONGEN system [14]. This is an aid to the organic chemist, offering him the computer generation of a hypothesis-tree of structures under given constraints. It can also be used as a verifier of claims of new structures, as a proof-checker. As an exercise in advanced organic chemistry, graduate students were assigned the verification of a set of structures recorded in the recent literature. Many of the proofs were found to be incomplete, usually for lack of tacit stipulations that were still plausible in the immediate context. We have no firm statistics, but perhaps one 'proof' in ten contained a substantive fallacy, unnoticed by the author and reviewers, that invited a critical reexamination of the conclusion. This suggests that organic chemical analysis has already become too complex for the existing media, that a significant part of the literature is shaky, and that computer-augmented proof-checking of complex structures should be part of the process of editorial review. The prevalence of statistical fallacies in the biomedical literature, often deeply rooted in careless experimental designs, has provoked much critical comment [15-18]. Certainly, it is responsible for a redoubled waste of resources, in the primary efforts, in faulty policy and practice, and in the further work needed for criticism and rectification.

Probably it is wrong to say that chemistry is so complex; to the contrary this finding is more likely a result of the simplicity and transparency of the logical argument in its proofs, which makes them more amenable to computer emulation. Outside of mathematics, very little scientific reasoning has been subjected to formal analysis and representation. EUGRAM publication now affords the opportunities and incentives to undertake more rigorous formulations both by providing more convenient media for depositing illegible proofs and offering access to symbol-manipulating machines to digest them. Increasingly, hardware engineers will find themselves companions to linguists and philosophers of science [19,20]; they have long since shared profitable joint ventures with formal logicians.

Emergence of the New Literacy [8,21]

The previous discussion declaims how the EUGRAM is a return to literacy, with some new forms and tools. The ease of its alteration saves some kinship of the EUGRAM to the oral tradition, with perhaps less social discipline but more effective tools to ensure the authenticity of the text. In fact, so much 'writing' is produced these days by dictation, with the most meagre and clumsy post-editing, that these tools may help bring the author closer to the welltempered text he intends. Most tools are two-edged: the ease of inserting cliches and of conforming to system-defined formats may also hinder creativity. But this is like agonizing whether desk calculators will frustrate arithmetic skills. Some authors will balk at learning to type -- even with all the facility of error correction afforded by every editor program. They can doubtless look forward within the decade to voice entry of rough texts that can speed up initial composition, and still leave scope for detailed editing. The author who does not interface directly with his own words with a text-display and editor is missing a powerful and precise organ of expression, which has no practical parallel in human communication today. Still, we can hardly surpass our inherent skills, though the wider availability of these compositional tools and challenges in education might help reverse the trend to illiteracy suggested by all recent statistics.

Not every communication will or should be reduced to an unerasable EUGRAM. Lovers will not be deterred, even by the black box, no more than they are by the mails; but other intimate communications -- particularly some of the angrier ones-- are better left to media where expletives can be deleted in hindsight. Even in scientific communication, there may be a place for a potential refuge: "I never said that?" in retrospection, namely to encourage some irresponsible imagination. This opportunity may be vitiated by the relentless accuracy of the EUGRAM, supported by new methods of encoding 'signatures'. Illegible handwritten scrawls will no longer offer a refuge of ambiguity. Nevertheless, while inscribed promises have more standing in court, voice-to-voice confrontation is less amenable to evasion at the moment: the journal-editor will telephone a delinquent author when repeated pleas by EUGRAM have been ignored. Conversely, the poetic imagination may be less hindered in a literate medium than in immediate confrontation with other critical voices. Ambiguous phrases can be left in the record, when they would be challenged in the vocal stream. These very assertions are ones that might be difficult to articulate in a lecture: they reveal mostly how little we know of the uses of different media.

Most of these remarks have concerned EUGRAMs between identified persons. The use of EUGRAMs for communication with archives opens up additional opportunities and foreseeable problems. In our experience at SUMEX-AIM, EUGRAPHY has been indispensable for the division of labor in drafting and criticizing complicated research proposals: 20 people may be closely involved in a product of 250 print pages. We have not secured a good system for tracking and interleaving successive versions, reducing a hairy tree of separate modifications to a coherent final form. Most nearly fatal is a cleanup reformatting that frustrates any simple line-by-line text comparison of deviant versions!

Confusions of this kind in communal refinement of encyclopedic texts can perhaps be ameliorated with further software for documentation control. However, they reflect an underlying difference between EUGRAMs, manuscripts, and unit copying on the one hand, and letterpress on the other. Gutenberg's method lodges the major cost of publication in composing a definitive version of a master template. A side effect of the economic advantage is the focus on that version as a node of the intellectual commitment of the author, and of criticism by others. Communal revision over a EUGRAM network is likely to outpace the reaction time of individual critics: Scientist "A" will be entering his critique of Heisenstein's Field Theory version 1764 when this has already been revised under the influence of "B" and superseded long since by version 1769. The same fluidity of commitment may be self-aggravating if scientists are then unconstrained in what they enter into the archives, believing that their errors are erasable, and that they must compete for priority with less scrupulous colleagues. The blurring of nodes of publication will also greatly complicate the task of assigning due credit for intellectual innovation, although in principle there can be greater technological support (auxiliary files and the like) for documenting the participation of many minds. The advantage of this fluidity is, obviously, a possible mitigation of prejudice and rigidity of beliefs that may otherwise impede intellectual progress.

The cost of nodal entry into letterpress also bolsters the gatekeeping role of editors and reviewers as trustees of the social investment entailed in that form of publication. This has already been eroded by the multiplication of commercial interests in scientific journals who receive an large unacknowledged subsidy a) in the public funding of the underlying research, and b) in the asset of attention of the readership. Both of these have been exploited to the point that existing publication is fragmented to an untolerated degree: namely, in many fields scientists no longer accept the responsibility for awareness of every claim that has reached print, particularly if these have bypassed the recognized, peer-refereed organs of their discipline. Near-zero-cost entry into the archives of a EUGRAM system will aggravate that problem, but has the compensation of an easy technology for selective retrieval. The role of the trustees will be shifted from controlling what enters the archives to that of organized consultation about what is worth perusing. Controversial innovations may be more fairly evaluated if minority approval is enough to permit them to reach the visible record.

The same technology can also be used to broaden the participatory base, and to reduce the grievous time lags and enhance the limited information flow that now characterizes peer review of research proposals used for the allocation of budgetary resources. The pros and cons of a wider base of 'voting' on one's colleagues' efforts can be roughly anticipated: in some sense more equitable distributions on the one hand; on the other, the factionalization of decision-making, political alliances, and the tyranny of the majority even in the most creative of individual activities. These dilemmas face us today; the new technologies will introduce a change of scale not of principle in the social monitoring of private thought. It is not just Big Brother we may need to fear, but the whole brood of our competing siblings.

The enemy may also be within ourselves. Scientists generally are systematically socialized within the norms of their profession; nevertheless they must approach the raging floods of literature with some ambivalence [22]: there might be found the nuggets of insight that may help the investigator take a bold new step. There is also the fear of finding an anticipation that may destroy the novelty, and hence the entire utility, of months, years or decades of sweat and the pride of unique intellectual accomplishment. The designer of information systems can ill afford to overlook Mooers' Law, that a "system will tend not to be used whenever it is more painful and troublesome for a customer not to have information than for him not to have it." [23]. Some writers tend to be egregiously neglectful of citing the roots of their ideas, a self-serving amnesia that also obscures others' access to the overall picture. The neglect also impedes the efficient retrieval of connected knowledge through devices like citation indexing. EUGRAM-based commentaries should facilitate the filling in of missing references by others, if the author has overlooked them, without making a major issue of the implied criticisms; and the anticipation of such corrections may deter the obliteration of the history of a subject. The cross-referencing and coding capabilities of bibliographic databases should also make it feasible for an author to exercise his historical responsibilities without excessively costly footnotes that may impede other uses of the entered material. In a similar vein, the systematic archiving of informal communications, including notes to oneself, surrounding the genesis of new ideas should facilitate the accurate reconstruction of the history of scientific discoveries -- narratives that today are inevitably clouded with more retrospective myth than documentable substance.

Altogether, we simply need to recognize that the new technology imposes fewer constraints per se, on the social structure of science, and that carefully designed new forms of social discipline will need to be established to meet the indicated functional needs.

The social innovations will doubtless evolve in response to microscopic pressures rather than as part of a system design, and their functionality will probably be tested on a time scale slower than continued technological inputs. Some of the needs and inventions can be foreseen; their main effect may be to facilitate another wave of illiteracy by the recruitment of still more elaborate devices for the human-bit interface. Reading and pecking are slow, and beneath the dignity of some professionals; voice response is even cheaper than the visual EUGRAM, and the technology for voice entry is on the way. Graphics already are an indispensable aid; there is no technological barrier to the integration of multi-modal cable-TV (e.g., animated cartoons) with EUGRAPHY. Programming costs will return the initiative to the centralized broadcaster; hopefully, a few individuals will still insist on their own selection of intellectual fare and many will sustain bilateral conversation. The literate tradition can still be enhanced with improved designs of orthographic display, a wider menu of formats including color, perhaps even new alphabets and languages. Indeed, it is language itself that needs more constructive as well as descriptive investigation: our existing tongues have evolved in response to long outmoded technologies of communication, but we know too little of the underlying neurobiology to be confident how they might be improved. Such studies are also impelled by the prevalence of pathologies of language development that constitute a heavy burden on many children and their schools. A 26-character alphabet certainly bears no relationship to any system that would be systematically designed to enhance the speed and reliability of human communications [24,25].

This discussion has intentionally focussed on the difficulties and side effects that may attend the introduction of challenging new technologies of communication [8,26]. Surely others will emerge as difficult to foresee as the impact of the internal combustion engine on the structure of cities. The problems should not obscure the constructive implications of steps towards the realization of an effective 'world brain', which had already obsessed Leibniz, and which may be the defining attribute of technological culture: the efficient refinement and sharing of human knowledge [27]. We do well to question our moral capability of enjoying the fruits of such cooperation; but this is not to damn ourselves in advance, especially if we acknowledge that anticipating the human problems is a task of equal priority to engineering the hardware.

Acknowledgment

The work at Stanford University, on which this essay is based, has been supported by the Biotechnology Resources Program, Division of Research Resources, National Institutes of Health (RR-00785; RR-00612) and by the Defense Advanced Research Projects Agency (Heuristic Programming Project, directed by Professor E. A. Feigenbaum). Many more people have contributed to the ideas presented here -- by publication, by personal interaction, by telephone, and by EUGRAPHY -- than I have been able to acknowledge in the available time and space.

References

[1] F.F. Clasquin and J.B. Cohen, "Prices of Physics and Chemistry Journals," Science, vol. 197, pp. 432-438, 29 July 1977.

[2] President's Science Advisory Committee, "Science, Government and Information," Washington: U.S.G.P.O., Jan. 10, 1963

[3] E. C. Anderson, J. B. Arnold, C. Emiliani, W. H. Johnston, R. L. Miller, H. E. Suess, and V. L. Telegdi, "Back to the Homeric Tradition." American Scientist, 1977, 43:117-119.

[4] J.R. Cole and S. Cole, "Social Stratification in Science," Chicago: Univ. Chicago Press, 1973.

[5] D. Crane, "Invisible Colleges," Chicago: Univ. Chicago Press, 1972.

[6] W. van den Daele and P. Weingart, "Resistance and Receptivity of Science to External Direction: The Emergence of New Disciplines under the Impact of Science Policy," pp. 247-275 in G. Lemaine et al.,eds., Perspectives on the Emergence of Scientific Disciplines, Chicago: Aldine, 1976.

[7] I. de Sola Pool, Ed., "The Social Impact of the Telephone," Cambridge, Mass.: The MIT Press, 1977.

[8] I. de Sola Pool, F. W. Frey, W. Schramm, N. Maccoby, and E. B. Parker, Eds., "Handbook of Communication," Chicago: Rand McNally, 1973.

[9] J.C.R. Licklider, "Libraries of the Future," Cambridge: M.I.T. Press, 1966.

[10] M. Turoff and S. R. Hiltz. "Meeting Through your Computer: Information Exchange and Engineering Decision-making are Made Easy Through Computer-Assisted Conferencing," IEEE Spectrum, May 1977, 58-64.

[11] G.W. Arnold and S.H. Unger, "A Structured Data Base Computer Conferencing System," National Computer Conference, pp. 461-467, 1977.

[12] R.Horton and R. Finnegan, "Modes of Thought," London: Faber and Faber,1973.

[13] K. Appel and W. Haken, "The Solution of the Four-Color Map Problem", in Scientific American, Oct. 1977.

[14] R.E. Carhart, S.M. Johnson, D.H. Smith, B.G. Buchanan, R.G. Dromey and J. Lederberg, "Networking and a Collaborative Research Community: A Case Study Using the DENDRAL Programs," pp. 192-217 in P. Lykos, ed., Computer Networking and Chemistry, Amer. Chem. Soc. Symposium Series No. 19, 1975.

[15] N.D.W. Lionel and A. Herxheimer, "Assessing Reports of Therapeutic Trials," Brit. Med. J., No. 3, pp. 637-640, 1970.

[16] S. Schor and I. Karten, "Statistical Evaluation of Medical Journal Manuscripts," J. Amer. Med. Assn., vol. 195, pp. 1123-1128, 1966.

[17] T. Dalenius, "Bibliography on Non-Sampling Errors in Surveys," Internl. Statis. Rev., vol. 45, pp. 71-89, 181-197, 303-317, 1977.

[18] I.J. Good, "Statistical Fallacies," in International Encyclopedia Social Sciences, New York: Macmillan, 1968.

[19] C.W. Churchman and B.G. Buchanan, "On the Design of Inductive Systems: Some Philosophical Problems", British Journal for the Philosophy of Science, vol. 20, pp.311-323, (1969).

[20] J. McCarthy, "Epistemological Problems of Artificial Intelligence" in Proceedings of the 5th International Joint Conference on Artificial Intelligence, MIT (Cambridge, Mass.), 1977.

[21] J. M. Nilles and F. R. Carlson, Jr., P. Gray, G. J. Hannenman, "The Telecommunications-Transportation Tradeoff: Options for Tomorrow," New York: John Wiley & Sons, 1976.

[22] R. K. Merton, "Sociological Ambivalence," New York: Free Press, 1976.

[23] C.N. Mooers, "Mooers' Law, or Why Some Retrieval Systems are used and Others are Not," American Documentation, vol 11., p. ii, July 1960.

[24] C. Cherry, "On Human Communication," New York: Wiley, 1957.

[25] W.A. Rosenblith, "Sensory Communication," Cambridge: M.I.T. Press, 1961.

[26] G. Gerbner, L. P. Gross, W. H. Melody, Eds., "Communications Technology and Social Policy," New York: John Wiley & Sons, 1973.

[27] M. Kochen, ed., "Information for Action: From Knowledge to Wisdom", New York: Academic Press, 1975.

Appendix III

COMPARISON OF MAINSAIL AND PASCAL

Clark R. Wilcox
Stanford University

MAINSAIL and PASCAL have been developed for different reasons, and it is this which is responsible for the major distinctions between them. The development of PASCAL was based on two principal aims, as stated by Wirth in the PASCAL USER MANUAL AND REPORT (henceforth referred to as the PASCAL REPORT):

> "The first is to make available a language suitable to teach programming as a systematic discipline based on certain fundamental concepts clearly and naturally reflected by the language. The second is to develop implementations of this language which are both reliable and efficient on presently available computers."

The basic goal of MAINSAIL, on the other hand, is to provide a machine-independent programming system suitable for the development of large, portable programs. PASCAL is a sparse, relatively simple language which is really more of a language kernel than a complete programming system. MAINSAIL is broader in scope, requires more runtime support and hence a more powerful processor, but does more for the programmer.

PASCAL as described in the PASCAL REPORT must be characterized as more of a blueprint for a language than a programming system. There are no compiletime facilities, very little standard runtime support, no standard access to a file system, and no concept of module. This lack of completeness, plus the elegance of design of what IS in PASCAL, is the reason for the proliferation of PASCAL implementations (no two of which are identical).

Of course there is no reason why an extended and portable version of PASCAL could not be created, and there has been some work in this area. But this was not Wirth's original goal, has not occurred in practice, and is not the language with which we are comparing MAINSAIL. Such a portable version would presumably be a more complete language, and hence would be a PASCAL-derivative (of which there are many) rather than PASCAL itself.

Portability

MAINSAIL is designed to support portable programs, and as a consequence the compiler and runtime system are largely written in MAINSAIL. The facilities provided by the portable compiler and runtime system are an inherent part of the language. In this sense MAINSAIL is more of a portable programming system than simply a programming language which can be implemented on many machines. A single compiler and runtime system which are used at all sites appears to be the only realistic means of obtaining the goal of portability. A language description such as that given in the PASCAL REPORT has never been sufficient, and there is no reason to believe that it ever will be.

Of course the most important consequence of portability is that programs designed with portability considerations in mind can be moved among implementations without alteration. A concerted effort has been made to guarantee the characteristics of a sufficiently rich programming environment that the programmer will seldom if ever feel the need to utilize machine-dependencies other than those which are inherent to the task being programmed.

Modules

Perhaps consistent with PASCAL's conception as a simple language, it has no concept of module, i.e., a program is a single unit which results from a single compilation.

To preserve machine-independence, MAINSAIL contains its own notion of inter-module communication, i.e., there is no reliance on a machine-dependent linkage system. MAINSAIL programs consist of independently compiled modules which may be executed from any address within memory, i.e., the modules are position-independent. Modules play a dual role as the vehicle for conceptual program modularization, and as the unit which is moved in and out of memory during execution to provide a virtual memory facility.

This precludes combining MAINSAIL modules with program fragments written in some other language, as could perhaps be done with some PASCAL implementations. However, MAINSAIL does provide for embedded assembly language code.

MAINSAIL encourages the view of a program as an open-ended collection of modules whose identity need not be known when the program is written. A flexible system of dynamic linkage allows arbitrary files which contain the executable code for a module to be read into memory and accessed from any other module which has declared the proper interface. Modules may be obtained from individual files, or from runtime libraries which are built and accessed via MAINSAIL system modules. Multiple instances of any module may coexist: an instance consists of shared code and a separate copy of data.

The lack of any such facilities in PASCAL must certainly be viewed as a limiting factor in its scope of applicability. Any but the simplest PASCAL implementation for machines with a small address space must deal with this deficiency, for otherwise large programs could not fit into memory.

Data types

PASCAL provides the standard data types boolean, integer, real, char, and pointer (defined in terms of another type). Perhaps PASCAL's most important contribution to programming languages is its simple yet extremely useful concepts of enumerated scalar types, subrange types, and set types. Operators for union, intersection, set difference, (in)equality, set inclusion and set membership are provided.

MAINSAIL provides the data types boolean, integer, long integer, real, long real, bits, long bits, string, pointer, address and charadr (character address). The latter two are so-called low-level types described in a later section. Bits and long bits provide 16- and 32-bit vectors which may take part in bitwise operations such as masking and shifting. Strings are described in the next section.

The ranges of integers, reals and chars in PASCAL are implementation-dependent. In MAINSAIL, integers are guaranteed the range provided by a 16-bit word, long integers and reals that provided by a 32-bit word, and long reals that provided by a 48-bit word. Thus the programmer knows what can be counted on, regardless of what machine executes the program.

MAINSAIL has none of PASCAL's user-declared type mechanism. The best one could do to "simulate" such types would be to use MAINSAIL's macro facility to get the effect of scalar types, use integers for scalar variables, and use the data type BITS to get the effect of sets. This is less readable than PASCAL's approach, and does not provide compiletime checking.

There are some draw-backs to PASCAL'S notion of types, but as usual these can be remedied by a more complete specification. The maximum size of a set is implementation-dependent. Also, PASCAL does not provide for input or output of scalar types or sets except via conversion to and from integer.

Strings

PASCAL has no string data type. Instead, it provides the type char (character). Packed char arrays provide fixed-length strings. Assignment of one packed char array to another involves copying all the characters. String constarts (sequences of chars enclosed in single quotes) can be assigned to packed char arrays. The programmer must keep track of the length of a string.

MAINSAIL provides a full implementation of variable-length strings. A string variable is implemented as a string descriptor which specifies the current length (number of characters) and the location of the first character. The characters are stored in a memory area called "string space". Whenever string space becomes full, MAINSAIL automatically compacts it by reclaiming characters which are no longer referenced by string descriptors.

In MAINSAIL, strings may be assigned, compared, concatenated, "substringed" and scanned in various manners. Other examples of system procedures for operating on strings are those for obtaining the length, first or last character; removing the first or last character; appending a character onto the front or end of a string; converting a value to its string representation and vice versa; reading and writing strings (as well as individual characters) from and to files; reading values from a string (as if reading from a text file); and writing values to a string (as if writing to a text file).

Arrays

PASCAL arrays are restrictive in two ways: they must have constant bounds, and they are statically allocated (unless a component of a dynamically allocated record). Aside from the obvious drawbacks of constant bounds, this restriction also has the unfortunate effect that all array arguments to a procedure with an array parameter must have the same constant bounds. For example, it is not possible to write a general-purpose sorting procedure which works on arrays of different bounds.

PASCAL has the concept of PACKED arrays, and the related procedures PACK and UNPACK to convert among packed and unpacked arrays. Packed arrays are presumably stored in a more compact form than usual, e.g., they can take advantage of subrange types to utilize the minimum bits per element.

PASCAL supports array assignment as a full copy of one array to another. Array comparison is allowed only for packed char arrays, which is PASCAL's representation for character strings.

MAINSAIL's arrays may have variable bounds, and their allocation and disposal is completely under user control. An array is implemented as a pointer to an array descriptor, which is a record which gives information necessary to access the array. The array storage itself is in a separate "record" which is referenced from the array descriptor. Array parameter passing, assignment, and comparison involve just the pointer to the array descriptor.

MAINSAIL's Init statement initializes an array with constant values. PASCAL provides no means of initializing an array other than assignment statements.

There are two penalties for MAINSAIL's more flexible notion of arrays. First, the array descriptor, which must be allocated along with the array storage, takes up storage. For small arrays (10 to 20 elements), the array descriptor is almost as big as the array itself, and is thus a significant overhead. Second, the extra indirection through the array descriptor commonly costs an extra instruction per element access.

Records

PASCAL has records as declared objects, as well as records allocated during execution and manipulated via a pointer. MAINSAIL has only the latter, in accordance with its design philosophy of no static addresses. Similarly, PASCAL has arrays of records and arrays of pointers, while MAINSAIL has only the latter. PASCAL's records must be explicitly disposed by the programmer, whereas MAINSAIL provides both explicit disposal and automatic "garbage collection".

PASCAL has record assignment which involves copying all the fields of the record, whereas MAINSAIL has a copy procedure for this purpose. In practice it is quite rare to copy a MAINSAIL record since it is usually just the pointers that are manipulated. However, PASCAL's static records require copying since they cannot be manipulated via pointers.

PASCAL's variant records and MAINSAIL's prefix classes serve a similar role in that both deal with record types (in MAINSAIL terminology, classes) which, though not identical, share some common fields. In PASCAL a single record type is declared which contains the common fields followed by a form of case selection to choose among the remaining "variant" fields. Whenever a record is created, a type constant is given which corresponds to the "tag field" (one of the common fields). This value is used to indicate which variant of the record to create.

In MAINSAIL, the common fields make up the fields of a separate class, say c. Separate classes are declared for each of the variant forms. These classes specify class c as a prefix class, which means that they inherent c's fields as their initial fields. Thus they all have the same initial fields, and the compiler is aware of this. Where PASCAL utilizes a single record declaration which incorporates the variants, MAINSAIL utilizes multiple class declarations with a common prefix class.

MAINSAIL pointer declarations can be more specific than PASCAL's with regard to variant records since in MAINSAIL a pointer is declared as referring to a particular class (which corresponds to a PASCAL variant), whereas in PASCAL a pointer can only be declared as referring to the record as a whole rather than a particular variant. The result is that MAINSAIL can catch some errors during compilation which are not detected by PASCAL.

MAINSAIL allows "unclassified" pointers which do not specify the class of records which they will point to. Such a pointer can point to any class, and thus the compiletime checking is not possible. This form is used as an escape for those situations for which classified pointers are too restrictive. Since PASCAL pointer declarations necessarily involve a type name, there is no way to deal with unclassified pointers.

PASCAL's WITH statement allows one or more pointers to be specified as default pointers over the scope of a statement. Record fields within the statement may be specified without being qualified by a pointer variable as long as one of the default pointers could be used with the field. The default pointers may not be modified within the statement. MAINSAIL has omitted such a facility since the statement interpretation becomes context dependent, the use of variable names the same as field names is restricted, and it is difficult for the compiler to enforce the rule that the default pointers cannot be modified.

Expressions

Unlike PASCAL, MAINSAIL has an If expression. MAINSAIL also provides an Assignment expression which allows the assignment operator to be used in expressions. MAINSAIL allows comparison chains such as "a < b < c" as an abbreviation of what must be used in PASCAL: "(a < b) AND (b < c)".

MAINSAIL's "dotted operators" are an extremely handy abbreviation:

a .op b	is an abbreviation for	a _ a op b
.- a	is an abbreviation for	a _ - a

where "op" is a binary operator such as "+".

PASCAL does not define the order of evaluation of the operands of AND and OR, whereas MAINSAIL guarantees that only as many operands are evaluated as are needed to determine the result. This is a great convenience; for example it is common to want to evaluate b only if a is TRUE in "a AND b", i.e., a serves as a "guard" on b. This is particularly useful in cases such as "WHILE p AND p.link DO ..." which is not so simply written in PASCAL.

Statements

MAINSAIL has an Expression statement which is simply a dotted expression as described earlier. Examples are

MAINSAIL	PASCAL
i .+ 1;	i := i + 1;
s .& "abc";	PASCAL has no strings
b .IOR procBit;	put an element into a set
a[i,j] .* 18;	a[i,j] := a[i,j] * 18;
.- k;	k := - k;

The Case statement is similar in both languages, except in PASCAL a scalar type can be used for case selection, while in MAINSAIL an integer must be used. MAINSAIL has the additional capabilities of allowing a case selector to specify a range of values, and to specify a default statement to be executed in the event that no case selector is satisfied. In MAINSAIL an error occurs if no selector is satisfied (and there is no catchall case); in PASCAL the result is undefined.

MAINSAIL has twelve forms of repetitive statements, whereas PASCAL has four. MAINSAIL provides a DONE statement to terminate an iteration, and a CONTINUE statement to continue an iteration. Both of these can be applied to any level from within a nested iteration. PASCAL provides no such facilities other than an unstructured Goto statement. PASCAL's lack of an iteration terminator causes either a redundant statement, an awkward use of Boolean variables, or a Goto statement.

MAINSAIL provides an explicit procedure return, whereas PASCAL does not. Instead a PASCAL function has an implicit variable given by the name and type of the function. At the end of execution the value of this variable is returned as the result. The lack of an explicit Return statement doubtless leads to a reliance on the Goto statement to get to the end of a function.

The Done, Continue and Return statements remove the need for a Goto statement in MAINSAIL. PASCAL's Goto statement utilizes numeric labels, a rather odd choice, especially considering that labels must be declared. The PASCAL REPORT does not define the effect of jumping into a structured statement.

Procedures

PASCAL makes a distinction between procedures, which do not return a value, and functions, which do return a value. MAINSAIL minimizes this distinction by referring to functions as typed procedures, and allowing typed procedures to be used in a statement (the result is simply discarded).

PASCAL has value and VAR (reference) parameters. A VAR parameter refers indirectly to the argument variable, and hence is subject to the well known confusions which can arise with this mechanism.

MAINSAIL has three parameter-passing mechanisms. A USES parameter is passed the value of its argument. A PRODUCES parameter is not initialized by its argument; instead it returns its value to the argument upon return from the procedure. A MODIFIES parameter combines the effect of a USES and a PRODUCES parameter: it is initialized by the argument, and it returns its value to the argument upon return from the procedure.

MAINSAIL provides OPTIONAL parameters whose arguments may be omitted in a procedure call (in which case zero of the proper data type is passed), and REPEATABLE parameters which may be passed multiple arguments (in which case the procedure is called multiple times, each time with the next repeated argument).

MAINSAIL's GENERIC procedures are another compiletime feature which provide a simple yet powerful means of using a generic procedure identifier in a call to represent any one of several different procedures as distinguished by the parameter types.

Procedures may be nested in PASCAL, but not in MAINSAIL. The division of a MAINSAIL program into relatively small modules, each of which contains relatively small procedures, virtually eliminates the need for further nesting of procedures.

Unlike PASCAL, MAINSAIL supports OWN variables, i.e., variables local to a procedure which retain their value over procedure entry and exit. PASCAL has parametric procedures and functions, while MAINSAIL does not.

File System

PASCAL says nothing about a file system. Instead there is a standard input file, a standard output file, local files and external files. The PASCAL REPORT states:

> "Files may be local to a program (or local to a procedure), or they may already exist outside the program. The latter are called external files. External files are passed as parameters in the program heading into the program."

Presumably the external files are set up by some means outside of PASCAL, but no mention is given of this mechanism. It is left as implementation-dependent, but even then there is no way provided for the program to have any control over the association of external files with file variables.

PASCAL has a file data type, but it is rather restrictive. A file is modeled as a sequence of components, all of the same type. The components can be accessed only sequentially.

Text files, i.e., files declared as type "FILE OF CHAR", require some extra mechanism. Since PASCAL does not define what characters terminate a line, special procedures writeln(f) and readln(f) are provided which write end-of-line or read up to an end-of-line, and the boolean procedure eoln(f) is true when the end of the current line has been reached in file f. Read and write are extended to allow reading and writing of integers and reals from text files (instead of just chars, which are the components of text files).

MAINSAIL considers a file to be a collection of data, on some external medium, that is treated as a unit by the file system (which is not a part of MAINSAIL). Files exist independently of the execution of a program, so that a program can create a file and associate it with a name which can later be used by another program to access the file. Thus unlike PASCAL, files can provide continuity from one program execution to another.

MAINSAIL makes a distinction between two file types: text and data. A text file consists of characters, and a data file consists of any mixture of numeric and bits data. MAINSAIL also distinguishes two methods of access to a file: sequential and random.

A file is referenced in a MAINSAIL program via a pointer that is produced by the file opening procedure. The pointer belongs to one of the classes "textFile" or "dataFile" which are predeclared by MAINSAIL. tty is a name for referring to the user's terminal. ttyRead and ttyWrite are system procedures used for explicit communication with tty. tty may also be "opened" and treated just like any text file, so that it can be used anonymously by a program. In PASCAL communication with the terminal is presumably provided via the standard input and output files.

In MAINSAIL, a command file (cmdFile) and logging file (logFile) are utilized for standard input and output. Normally these are associated with tty, but they may be redirected to any text file by the programmer, for example to get the effect of a batch stream.

MAINSAIL uses the predefined (implementation-dependent) string constant eol (end-of-line) as a line terminator, instead of special procedures such as PASCAL's readln and writeln. MAINSAIL also defines the string constant eop (end-of-page) as a page terminator. This use of characters to delimit lines and pages is more flexible than PASCAL's use of special procedures since it allows these indicators to be part of a string, and hence implicitly manipulated as data.

Compiletime Facilities

MAINSAIL has a comprehensive set of compiletime facilities which are invaluable in the construction of large programs. The only related facility in PASCAL is the ability to declare constants, and the closely related concept of scalar type declarations (a scalar type can be viewed as a structured set of constant declarations). PASCAL's lack of compile-time facilities reflects its conception as a simple one-module language rather than a tool for building large programs. The following is a summary of MAINSAIL's compiletime features.

MAINSAIL provides compiletime evaluation for constant expressions, i.e. expressions involving only constant operands. A full macro facility provides definition of constants and arbitrary text with optional macro parameters. The programmer can interactively define macros during compilation.

The Message directive directs the compiler to print a message during compilation. The Sourcefile directive specifies a file which is to be compiled as if its text appeared in place of the directive. Thus text which is to be used in several modules, such as a "macro library", can be placed in a single file and sourcefiled from all the modules.

Conditional compilation allows the programmer to specify under what conditions indicated parts of the source file are to be compiled or ignored. Scanning directives allow pages in the source file to be skipped, and provide for explicit termination of compilation of the current file as if the end of the file had occurred.

A facility is provided for automatic utilization of compiletime libraries, which are just files which contain procedure bodies which are to be "compiled into" a number of different modules.

Save and restore directives allow the compiler's symbol table to be saved, and then restored during some other compilation. This avoids recompilation of frequently used "header" files such as macro libraries.

Other directives give the programmer control over the amount of code emitted to check error conditions such as array subscripts out of bounds and null pointers used for field access.

Low-level Features

MAINSAIL provides features which allow a low-level access to the host machine. They are for use only by knowledgeable programmers who need access to underlying representations or who need to intersperse some machine-dependent code with MAINSAIL code. These features are extensively used by the MAINSAIL runtime system, and thus are an integral part of the language. Nevertheless, MAINSAIL provides enough facilities that those described here can usually be avoided.

The data type ADDRESS is provided for representing arbitrary memory addresses. To access individual characters, the data type CHARADR (character address) is provided. A number of supporting features are included which allow addresses and charadrs to be used for access to unstructured memory.

The Code statement allows assembly language to be included as a string constant which is simply put into the compiler's output file. The ENCODE directive and CODED procedures provide additional capabilities for assembly language programming. Of course, modules which contain assembly language are machine-dependent.

MAINSAIL's low-level features allow procedures to be written in MAINSAIL which would otherwise have to be coded in assembly language. They allow almost all of MAINSAIL's runtime system to be written in MAINSAIL (even the machine-dependent parts). Even those parts which must be written in assembly language (e.g., monitor calls) may be included in MAINSAIL modules. PASCAL has no such features, so that it cannot express manipulation of arbitrary memory addresses.

Appendix IV

AIM MANAGEMENT COMMITTEE MEMBERSHIP

The following are the membership lists of the various SUMEX-AIM management committees at the present time:

AIM EXECUTIVE COMMITTEE:

LEDERBERG, Joshua, Ph.D. (Chairman)
Department of Genetics, S331 [through 8/78]
Stanford University Medical Center
Stanford, California 94305
(415) 497-5801

The Rockefeller University [after 9/78]
1230 York Avenue
New York, New York 10021
(212) 360-1234

AMAREL, Saul, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546

BAKER, William R., Jr., Ph.D. (Executive Secretary)
Biotechnology Resources Program
National Institutes of Health
Building 31, Room 5B43
9000 Rockville Pike
Bethesda, Maryland 20014
(301) 496-5411

COHEN, Stanley N., M.D.
Department of Genetics and
Division of Clinical Pharmacology,
Department of Medicine, S155
Stanford University Medical Center
Stanford, California 94305
(415) 497-5315

FEIGENBAUM, Edward, Ph.D.
Department of Computer Science
Polya Hall, Room 213
Stanford University
Stanford, California 94305
(415) 497-4079

LINDBERG, Donald, M.D. (Adv Grp Member)
605 Lewis Hall
University of Missouri
Columbia, Missouri 65201
(314) 882-6966

MYERS, Jack D., M.D.
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 624-2649

AIM ADVISORY GROUP:

LINDBERG, Donald, M.D. (Chairman)
605 Lewis Hall
University of Missouri
Columbia, Missouri 65201
(314) 882-6966

AMAREL, Saul, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546

BAKER, William R., Jr., Ph.D. (Executive Secretary)
Biotechnology Resources Program
National Institutes of Health
Building 31, Room 5B43
9000 Rockville Pike
Bethesda, Maryland 20014
(301) 496-5411

COHEN, Stanley N., M.D.
Department of Genetics and
Division of Clinical Pharmacology,
Department of Medicine, S155
Stanford University Medical Center
Stanford, California 94305
(415) 497-5315

FEIGENBAUM, Edward, Ph.D.
Department of Computer Science
Polya Hall, Room 213
Stanford University
Stanford, California 94305
(415) 497-4079

LEDERBERG, Joshua, Ph.D. (Ex-officio)
Principal Investigator - SUMEX
Department of Genetics, S331
Stanford University Medical Center
Stanford, California 94305
(415) 497-5801

MINSKY, Marvin, Ph.D.
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139

MOHLER, William C., M.D.
Associate Director
Division of Computer Research and Technology
National Institutes of Health
Building 12A, Room 3033
9000 Rockville Pike
Bethesda, Maryland 20014
(301) 496-1168

MYERS, Jack D., M.D.
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 624-2649

PAUKER, Stephen G., M.D.
Department of Medicine - Cardiology
Tufts New England Medical Center Hospital
171 Harrison Avenue
Boston, Massachusetts 02111
(617) 956-5910

SIMON, Herbert A., Ph.D.
Department of Psychology
Baker Hall, 339
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213
(412) 578-2787 or 578-2000

STANFORD COMMUNITY ADVISORY COMMITTEE:

LEDERBERG, Joshua, Ph.D. (Chairman)
Principal Investigator - SUMEX
Department of Genetics, S331
Stanford University Medical Center
Stanford, California 94305
(415) 497-5801

COHEN, Stanley N., M.D.
Department of Genetics and
Division of Clinical Pharmacology,
Department of Medicine, S155
Stanford University Medical Center
Stanford, California 94305
(415) 497-5315

DJERASSI, Carl, Ph.D.
Department of Chemistry, Stauffer I-106
Stanford University
Stanford, California 94305
(415) 497-2783

FEIGENBAUM, Edward, Ph.D.
Department of Computer Science
Polya Hall, Room 213
Stanford University
Stanford, California 94305
(415) 497-4079

LEVINTHAL, Elliott C., Ph.D.
Department of Genetics, S047
Stanford University Medical Center
Stanford, California 94305
(415) 497-5813