S U M E X

STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE

RR - 00785

ANNUAL REPORT - YEAR 11

Submitted to

BIOTECHNOLOGY RESOURCES PROGRAM
NATIONAL INSTITUTES OF HEALTH

June 1, 1984

STANFORD UNIVERSITY SCHOOL OF MEDICINE
Edward A. Feigenbaum, Principal Investigator
Edward H. Shortliffe, Co-Principal Investigator

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
NATIONAL INSTITUTES OF HEALTH

DIVISION OF RESEARCH RESOURCES
BIOTECHNOLOGY RESOURCES PROGRAM

ANNUAL PROGRESS REPORT
PART I., TITLE PAGE

1. PHS GRANT NUMBER: **5P41RR00785-11**

2. TITLE OF GRANT: **SUMEX**

   Stanford University Medical Experimental Computer Resource

3. NAME OF RECIPIENT INSTITUTION: **Stanford University**

4. HEALTH PROFESSIONAL SCHOOL: **School of Medicine**

5. REPORTING PERIOD:

   5a. FROM: **08-01-83**
   5b. TO: **07-31-84**

6. PRINCIPAL INVESTIGATOR:

   6a. NAME: **Edward A. Feigenbaum, Ph.D.**
   6b. TITLE: **Professor of Computer Science**

   6c. SIGNATURE: Edward A. Feigenbaum

7. DATE SIGNED: May 18, 1984

8. TELEPHONE: **415-497-4879**

Table of Contents

I. Narrative Description 1
I.A. Summary of Research Progress 2
I.A.1. Overview of Objectives and Rationale 2
I.A.1.1. What is Artificial Intelligence 2
I.A.1.2. Impact of AI in Biomedicine 4
I.A.2. Details of Technical Progress 7
I.A.2.1. Facility Management and Operation 7
I.A.2.2. Facility Management and Operation 7
I.A.2.3. Timesharing Systems 16
I.A.2.4. Professional Workstations 21
I.A.2.5. Networking and Communications 24
I.A.2.6. Progress in Core Research 28
I.A.2.7. Resource Operations Statistics 30
I.A.2.8. SUMEX Staff Publications 46
I.A.2.9. Future Plans 47
I.B. Highlights 53
I.B.1. Progress Towards a Distributed SUMEX-AIM 54
I.B.2. New Molgen Directions 55
I.B.3. ONCOCIN - An Oncology Chemotherapy Advisor 56
I.B.4. New Pilot Projects 57
I.B.5. Major Books on Medical Artificial Intelligence 59
I.B.6. Training in Medical Information Science 60
I.C. Administrative Changes 61
I.D. Resource Management and Allocation 62
I.D.1. Management Committees 62
I.D.2. New Project Recruiting 63
I.D.3. Stanford Community Building 64
I.D.4. Existing Project Reviews 64
I.D.5. Resource Allocation Policies 64
I.E. Dissemination Efforts 66
I.F. Comments on the Biotechnology Resources Program 68
II. Description of Scientific Subprojects 69
II.A. Scientific Subprojects 69
II.A.1. Stanford Projects 70
II.A.1.1. EXPEX - Expert Explanation Project 71
II.A.1.2. GUIDON/NEOMYCIN Project 78
II.A.1.3. HPP Core AI Research 89
II.A.1.4. MOLGEN Project 98

II.A.1.5. ONCOCIN Project 106
II.A.1.6. RADIX Project 121
II.A.2. National AIM Projects 129
II.A.2.1. CADUCEUS Project 130
II.A.2.2. CLIPR - Hierarchical Models of Human Cognition 136
II.A.2.3. Rutgers Research Resource 142
II.A.2.4. SECS: Simulation & Evaluation of Chemical Synthesis 149
II.A.2.5. SOLVER Project 159
II.A.3. Pilot Stanford Projects 166
II.A.3.1. CAMDA Project 167
II.A.3.2. MENTOR Project 175
II.A.3.3. Protein Secondary Structure Project 179
II.A.3.4. PROTEAN Project 183
II.A.3.5. Ultrasonic Imaging Project 185
II.A.4. Pilot AIM Projects 190
II.A.4.1. PATHFINDER Project 191
II.A.4.2. RXDX Project 197
II.B. Books, Papers, and Abstracts 202
II.C. Resource Summary Table 202
Appendix A. AIM Management Committee Membership 203
Appendix B. Community Growth and Project Abstracts 209

List of Figures

Figure 1:	Current SUMEX-AIM Decsystem 2060 Computer Configuration	10
Figure 2:	Current SUMEX-AIM 2020 Computer Configuration	11
Figure 3:	Current Shared VAX Computer Configuration	12
Figure 4:	SUMEX-AIM Ethernet Configuration	13
Figure 5:	SUMEX-AIM File Server {SAFE}	14
Figure 6:	SUMEX-AIM Development Vax {ARDVAX}	15
Figure 7:	Total CPU Time Consumed by Month	31
Figure 8:	Peak Number of Jobs by Month	32
Figure 9:	Peak Load Average by Month	32
Figure 10:	Monthly CPU Usage by Community	34
Figure 11:	Monthly Terminal Connect Time by Community	35
Figure 12:	Cumulative CPU Usage Histogram by Project and Community	37
Figure 13:	TYMNET Terminal Connect Time	45
Figure 14:	ARPANET Terminal Connect Time	45
Figure 15:	SUMEX-AIM Growth by Community	209

I. Narrative Description

This third year of the 5-year renewal of the SUMEX resource grant has been an active year not only for the SUMEX staff, but for the SUMEX-AIM community involved in developing expert systems. Successes in developing such systems, many of them stemming from projects in the SUMEX-AIM community, continue to stimulate strong and growing interest in AI research on many educational, governmental, and industrial fronts.

This is an annual report for the **S**tanford **U**niversity **M**edical **EX**perimental computer resource for applications of **A**rtificial **I**ntelligence in **M**edicine (**SUMEX-AIM**). It covers the period between May 1, 1983 and April 30, 1984.

This third year of the 5-year renewal of the SUMEX resource grant has been an active year not only for the SUMEX staff, but for the SUMEX-AIM community involved in developing expert systems. Successes in developing such systems, many of them stemming from projects in the SUMEX-AIM community, continue to stimulate strong and growing interest in AI research on many educational, governmental, and industrial fronts.

In addition, this past year has seen concurrent development of SUMEX-AIM as a distributed scientific resource. Our approved project goals focus principally on the merging of state-of-the-art community research in biomedical AI applications with new computing tools and on the challenges they will bring to the SUMEX-AIM community and resource. The SUMEX staff continues to exploit these advances in professional workstations and communication technology, while at the same time maintaining our high standards for a computing resource.

This third year also saw the initiation of a number of SUMEX-AIM pilot projects. These pilot projects provide new activities and research directions for the community to replace existing projects which have matured and moved off the SUMEX-AIM resource.

The earlier phases of the SUMEX-AIM resource were characterized by the building of a national community of biomedical AI collaborators around a central resource located at Stanford University. Beginning with 5 projects in 1973, the AIM community grew to 11 major projects at our renewal in 1978. This past year saw the completion of two long term and successful projects on SUMEX-AIM; DENDRAL and PUFF/VM. There currently are 13 fully-authorized projects plus seven pilot efforts.

Many of the computer programs under development by these groups are maturing into tools increasingly useful to the respective research or clinical communities. We continue to seek out new AI applications in our community of biomedical and computer scientists who interact through electronic media. The SUMEX-AIM community is beginning to evolve as a highly distributed resource, with the SUMEX staff and computer facility serving as the backbone to electronic communication and systems support. The community is becoming more and more involved in personal computers and professional workstations, and more heavily dependent on network communication facilities for interactions, collaborations, and sharing.

The following sections cover the activities of the SUMEX-AIM resource this past year, including brief summaries of the our objectives, a characterization of biomedical AI research, resource organization and operating procedures, recent core progress in system development and basic AI research, and progress in the collaborative projects.

I.A. Summary of Research Progress

I.A.1. Overview of Objectives and Rationale

SUMEX-AIM ("SUMEX") is a national computer resource with a dual mission: 1) promoting applications of computer science research in artificial intelligence (AI) to biological and medical problems, and 2) demonstrating computer resource sharing within a national community of health research projects. The central SUMEX-AIM facility is located physically 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.

I.A.1.1. What is Artificial Intelligence

The subfield of computer science known as Artificial Intelligence, or AI, deals with symbolic reasoning using large amounts of heuristic knowledge. Many of the world's difficult problems are symbolic, such as troubleshooting electronic or mechanical equipment, medical diagnosis and therapy planning, and configuring elemental parts into a whole system. For these kinds of problems, AI offers new opportunities for developing computer-based solutions.

In addition to using symbolic representations of knowledge, AI also uses heuristic methods for processing information. Heuristics are rules of thumb, judgmental rules that aid in finding ***plausible*** solutions. AI is distinguished from other areas of computing in its attention to both symbolic (non-numeric) information and heuristic (non-algorithmic) methods for solving problems.

Placing AI in Computer Science

The major focus of AI is understanding intelligence through construction (or programming) of machines that behave intelligently. That is a grand goal. In the short-term, AI research focuses on non-numerical problem solving in order to build experience with problem solving methods, techniques for representing various kinds of knowledge, interfaces with users, and numerous other issues.

One of the distinguishing features of problems for which AI methods have been developed is that the problems are not well-structured. That is, one does not already know in advance (from the problem description alone) what the best method is for solving the problem. In short, there are no algorithms. Broadly speaking, AI substitutes exploratory search for precise, algorithmic solution methods.

Expert Systems and Applications

The national SUMEX-AIM resource is an outgrowth of a long, interdisciplinary line of artificial intelligence research at Stanford and elsewhere concerned with the development of concepts and techniques for building "expert systems" [1]. An "expert system" is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. For some fields of work, the knowledge necessary to perform at such a level,

plus the inference procedures used, can be thought of as a model of the expertise of the expert practitioners of that field.

Two important features that distinguish expert systems from conventional programs are flexibility and understandability. Expert systems are ***flexible*** in the sense that they can be changed and extended easily, and they are ***understandable*** in the sense that they can explain the contents of their own knowledge bases and their own lines of reasoning [10]. These features are especially important in medicine, where knowledge is changing rapidly and where practitioners have to understand the reasons for a program's decisions because they have to accept responsibility for following (or not following) those decisions.

The application areas range from medicine to electronics, from machinery to software. The problems range from diagnosis and troubleshooting (analysis) problems to planning and configuration (synthesis) problems. Knowledge bases for expert systems are built iteratively -- usually through long interactions over many months between a human specialist who understands the details of the domain and a knowledge engineer who understands the programming details of the system.

The knowledge of an expert system consists of facts and heuristics. The "facts" constitute a body of information that is widely shared, publicly available, and generally agreed upon by experts in a field. The "heuristics" are the mostly-private, little-discussed rules of good judgment (rules of plausible reasoning, rules of good guessing) that characterize expert-level decision making in the field. The performance level of an expert system is primarily a function of the size and quality of the knowledge base that it possesses. One of the key ideas in maintaining flexibility and understandability is the clean separation of elements of the knowledge base from elements of the program that interpret the knowledge base.

The major issues in building expert systems, at the moment, are:

- selecting an appropriate problem (in terms of size, difficulty, importance, decomposability, risk)
- selecting a representation and control structure (or framework system that supplies both),
- settling on an appropriate vocabulary and conceptualization for the problem,
- finding an available expert,
- transferring the expert's knowledge into the program (knowledge engineering),
- refining the knowledge base with feedback from test cases,
- packaging the system in a form that is acceptable to end-users,
- validating the quality of the program's advice.

One of the best known expert systems is MYCIN [3], a program in which the separation of knowledge (of medicine) from the rest of the program was carefully engineered. (The abstracted case of an arbitrary knowledge base and a framework interpreter, plus auxiliary programs, was achieved in the EMYCIN system [16], to which knowledge of other domains can be added to build a diagnostic system in those domains.)

Currently authorized projects in the SUMEX community are concerned in some way with the application of AI to biomedical research*. The tangible objective of this approach is the development of computer programs that will be more general and effective

*Brief abstracts of the various projects can be found in Appendix B on page 209 and more detailed progress summaries in Section II on page 69.

consultative tools for the clinician and medical scientist. There already have been promising results in areas such as chemical structure elucidation and synthesis, diagnostic consultation, molecular biology, and modeling of psychological processes.

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 comprising artificial intelligence research. 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.

I.A.1.2. Impact of AI in Biomedicine

There is a certain inevitability to the field of Artificial Intelligence and its applications, in particular, to medicine and biosciences. The cost of computers will continue to fall drastically during the coming two decades. As it does, many more of the practitioners of the world's professions will be persuaded to turn to economical automatic information processing for assistance in managing the increasing complexity of their daily tasks. They will find, from most of computer science, help only for those problems that have a mathematical or statistical core, or are of a routine data-processing nature. But such problems will be relatively rare, except in engineering and physical science. In medicine, biology, management, indeed in most of the world's work, the daily tasks are those requiring symbolic reasoning with detailed professional knowledge. The computers that will act as ***intelligent assistants*** for these professionals must be endowed with symbolic reasoning capabilities and knowledge.

The growth in medical knowledge has far surpassed the ability of a single practitioner to master it all, and the computer's superior information processing capacity thereby offers a natural appeal. Furthermore, the reasoning processes of medical experts are poorly understood; attempts to model expert decision-making necessarily require a degree of introspection and a structured experimentation that may, in turn, improve the quality of the physician's own clinical decisions, making them more reproducible and defensible. New insights that result may also allow us more adequately to teach medical students and house staff the techniques for reaching good decisions, rather than merely to offer a collection of facts which they must independently learn to utilize coherently.

The knowledge that must be used is a combination of factual knowledge and heuristic knowledge. The latter is especially hard to obtain and represent since the experts providing it are mostly unaware of the heuristic knowledge they are using. Medical and scientific communities currently face many widely-recognized problems relating to the rapid accumulation of knowledge, for example:

- codifying theoretical and heuristic knowledge
- effectively using the wealth of information implicitly available from textbooks, journal articles and other practitioners
- disseminating that knowledge beyond the intellectual centers where it is collected
- customizing the presentation of that knowledge to individual practitioners as well as customizing the application of the information to individual cases

We believe that computers are an inevitable technology for helping to overcome

these problems. While recognizing the value of mathematical modeling, statistical classification, decision theory and other techniques, we believe that effective use of such methods depends on using them in conjunction with less formal knowledge, including contextual and strategic knowledge.

Artificial intelligence offers advantages for representing and using information that will allow physicians and scientists to use computers as intelligent assistants. In this way we envision a significant extension to the decision-making powers of specific practitioners without reducing the importance of those individuals in that process.

Knowledge is power, in the profession and in the intelligent agent. As we proceed to model expertise in medicine and its related sciences, we find that the power of our programs derives mainly from the knowledge that we are able to obtain from our collaborating practitioners, not from the sophistication of the inference processes we observe them using. Crucially, the knowledge that gives power is not merely the knowledge of the textbook, the lecture and the journal, but the knowledge of ***good practice***--the experiential knowledge of ***good judgment*** and ***good guessing***, the knowledge of the practitioner's art that is often used in lieu of facts and rigor. This heuristic knowledge is mostly private, even in the very public practice of science. It is almost never taught explicitly, is almost never discussed and critiqued among peers, and most often is not even in the moment-by-moment awareness of the practitioner.

Perhaps the the most expansive view of the significance of the work of the SUMEX-AIM community is that a methodology is emerging for the systematic explication, testing, dissemination, and teaching of the heuristic knowledge of medical practice and scientific performance. It may be less important that computer programs can be organized to use this knowledge than that the knowledge itself can be organized for the use of the human practitioners of today and tomorrow.

Evidence of the impact of SUMEX-AIM in promoting ideas such as these, and developing the pertinent specific techniques, has been the explosion of interest in medical artificial intelligence and the specific research efforts of the SUMEX community. As SUMEX has entered its second decade, we have found that the small community of researchers that characterized the AIM field in the early 1970's has now grown to a large, accomplished, and respected research community. The American Association for Artificial Intelligence (AAAI), the principal scientific membership organization for the AI field, has 4000 members, over 1000 of whom are members of the medical special interest group known as the AAAI-M. This subgroup was founded by members of the SUMEX-AIM community who were active in AAAI and is the only active subgroup in the Association. The organization distributes semiannual newsletters on medical AI and provides a focus for co-sponsoring relevant medical computing meetings with other societies (such as the American Association for Medical Systems and Informatics -- AAMSI). Medical AI papers are prominently featured at both medical computing and artificial intelligence meetings, and artificial intelligence is now routinely featured as a specific subtopic for specialized sessions at medical computing and other medical professional meetings. For example, members of the AIM community have represented the field to physicians at the American College of Pathology and American College of Physicians meetings for the last several years. A mere decade ago, the words "artificial intelligence" were never uttered at such conferences. The growing interest and recognition are largely due to the activities of the SUMEX-AIM community.

Another indication of the growing impact of the SUMEX-AIM community is its effect on medical education. For reasons such as those outlined above, there is an increasing recognition of the need for a revolution in the way medicine is taught and

medical students organize and access information. Computing technology is routinely cited as part of this revolution, and artificial intelligence (and SUMEX-AIM research) generally figures prominently in such discussions. Such diverse organizations as the National Library of Medicine, the American College of Physicians, the Association of American Medical Colleges, and the Medical Library Association have all called for sweeping changes in medical education, increased educational use of computing technology, enhanced research in medical computer science, and career development for people working at the interface between medicine and computing; reports of all four organizations have specifically cited the role of artificial intelligence techniques in future medical practice and have used SUMEX-AIM programs as examples of where the technology is gradually heading.

In summary, the logic which mandates that artificial intelligence play a key role in enhancing knowledge management and access for biomedicine -- a logic in which we have long believed -- has gradually become evident to much of the biomedical community. We are encouraged by this increased recognition, but realistic about the significant research challenges that remain. Our goals are accordingly both scientific and educational. We continue to pursue the research objectives that have always guided SUMEX-AIM, but must also undertake educational efforts designed to inform the biomedical community of our results while cautioning it about the challenges remaining.

I.A.2. Details of Technical Progress

I.A.2.1. Facility Management and Operation

The following material covers the SUMEX-AIM resource activities over the past year in greater detail. Individual sections cover progress in ;

- Facility Management and Operation
- Timesharing Systems
- Professional Workstations
- Networking and Communications

These sections outline accomplishments in the context of the resource staff and resource management. Details of the progress and plans for our external collaborative projects are presented in Section II beginning on page 69.

I.A.2.2. Facility Management and Operation

SUMEX-AIM continues to manage and operate it's computing resources in a effective and efficient manner conducive to providing a reliable and robust computing environment.

While the previous year (Year 10) involved a major move from the KI10 Tenex system to a new DECsystem 2060, this year saw more emphasis to our gradual move to distributed processing, while continuing to improve our excellent timesharing environment on the 2060. This development is covered in full in section I.A.2.2 starting on page 15.

Our continued movement to professional workstations has taken on several forms. We have continued to acquire Lisp machines for use by the SUMEX community while at the same time investigating the use of remote virtual graphics and new lower cost workstations such as the Apple Macintosh, Sun workstations, and others that are appearing on the market. The development of professional workstations is covered in more detail in section I.A.2.3 starting on page 21.

SUMEX continues to expend a great deal of effort in the support and development of our networking and communications facilities. Key to our ability to provide the maximum computing power available to the greatest number of users is a mechanism for making it irrelevant where that user is physically located. By having a robust networking and communications environment, we are able to extend our facility to any user or group of users, thereby making available to them the power and convenience of SUMEX. Further information on the progress made in networking and communications can be found in section I.A.2.4 starting on page 23.

In the area of facility management and operation, several noteworthy events occurred over the past year which will be explained in more detail here.

SUMEX/HPP Welch Road Computing Facility

A major development this past year at SUMEX was the move of the Heuristic Programming Project to their new location at 701 Welch Road, adjacent to the Stanford Medical Center. Since this group is a major user of the SUMEX-AIM resource and the focus for most of the core AI research, a good deal of effort was expended to provide a robust computing environment at their new location. This development involved several stages and levels of technical development, ranging from construction of the machine room, new cable and wiring installation, procurement and setup of networking hardware, to major new developments in the networking software and a twisted pair Ethernet communication link between this site and the main SUMEX Computer room. All of the hardware and facilities purchases were funded from sources other than SUMEX.

We setup the general communications capabilities for the two buildings occupied by the HPP. This involved wiring up local terminals, installing local Ethernets (both 3 and 10 megabit capability), and acquiring and installing networking hardware such as terminal interface processors (TIP) and gateways, as well as extending the current SUMEX TIP and GATEWAY software to handle both 3 and 10 megabit network traffic.

But the most important and most interesting development in this process was the "twisted pair' ethernet developed by the SUMEX engineering staff to allow high speed reliable communications between this Welch Road facility and the SUMEX machine room. Further information on this new ethernet can be found in Section I.A.2.4 on page 23.

HPP researchers are routinely using this link to communicate with SUMEX and the central university network. In addition, various Lisp machines and printers located in the HPP facility and connected to a local network are able to communicate with the university network.

The end result is that we have successfully been able to extend the SUMEX computing environment to a remote site, providing a high speed link to the facilities of SUMEX while also allowing for local distributed processing. We see this experience has being most valuable in the future as we move further into a distributed environment, while still needing the sharing of resources and communication links provided by large timesharing systems and local area networks.

Digital Equipment Corporation stops development of 36-bit product line

Digital Equipment Corporation, a long time supplier of high speed 36-bit timesharing computers to the Artificial Intelligence community, announced that it was stopping all development of future 36-bit products, and instead starting a program to provide a migration path to its line of VAX minicomputers.

Many DEC 20 customers had been anticipating a new yet unannounced machine from DEC code named the 'Jupiter', which had been reported to be a order of magnitude faster than the current KL10 processor used in DEC20's and DEC10's. However, DEC's announcement means this effort has stopped, and we can expect no more 36-bit products from Digital Equipment Corporation.

The effect of this announcement to the AI community is disappointing, although not totally unexpected. The DECsystem20 has been the predominant timesharing machine used to support Artificial Intelligence based research, but yet researchers have been in need of more processing power and larger address spaces for quite a few years. DEC has clearly decided to devote their resources to VAX development. For those in need of greater 36-bit processing power or address space, you must now look to newer less

experienced companies such as Foonley and Systems Concepts for follow on 36-bit products which those firms are preparing.

The impact of this decision on the SUMEX-AIM community must be examined in conjunction with our development of AI systems on personal Lisp machines. We have outlined very clearly our plans to move to distributed Lisp-based workstations for AI Research, and this is clearly where we see the AI computing market heading. These machines offer much better cost/performance ratios than timesharing machines, high resolution bit-mapped screens, and powerful Lisp programming environments for the development and eventual dissemination of AI based systems. However, this is not too say we still do not see a role for the large timesharing machine in our environment. We still believe in the use of a large central mainframe computer as the anchor for a large community of users. The mainframe also functions as a central facility for communication and collaboration, and provides fast Lisp cycles for program development where the application is not in need of a specialized workstation.

Other SUMEX Computing Facilities

SUMEX continues to support other mainframe computers, file servers, professional workstations, and assorted printers and terminals for use by the SUMEX-AIM community.

1. The SUMEX-AIM File server, based on a VAX 11/750 computer, continues to serve the needs of the workstation users within SUMEX-AIM. The use of SAFE by users of our 2060 is minimal. We plan to extend the use of SAFE in the future by providing more convenient access by 2060 users than is currently available.

2. The VAX 11/780 computer system, originally purchased with DARPA funds and previously located in Margaret Jacks Hall on campus, has become a SUMEX-AIM resource this past year. The system was moved to a new location on the Stanford campus which provides a better environment for a computer of this size. This VAX is now shared between the Computer Science Department and the SUMEX-AIM community.

3. SUMEX continues to support a wide range of professional workstations from such vendors as Xerox, Symbolics, and Hewlett Packard for the development and testing of AI applications. Additional work has been started to explore the use of the Apple Macintosh and Apple Lisa within SUMEX. More information on these developments can be found in section I.A.2.3.

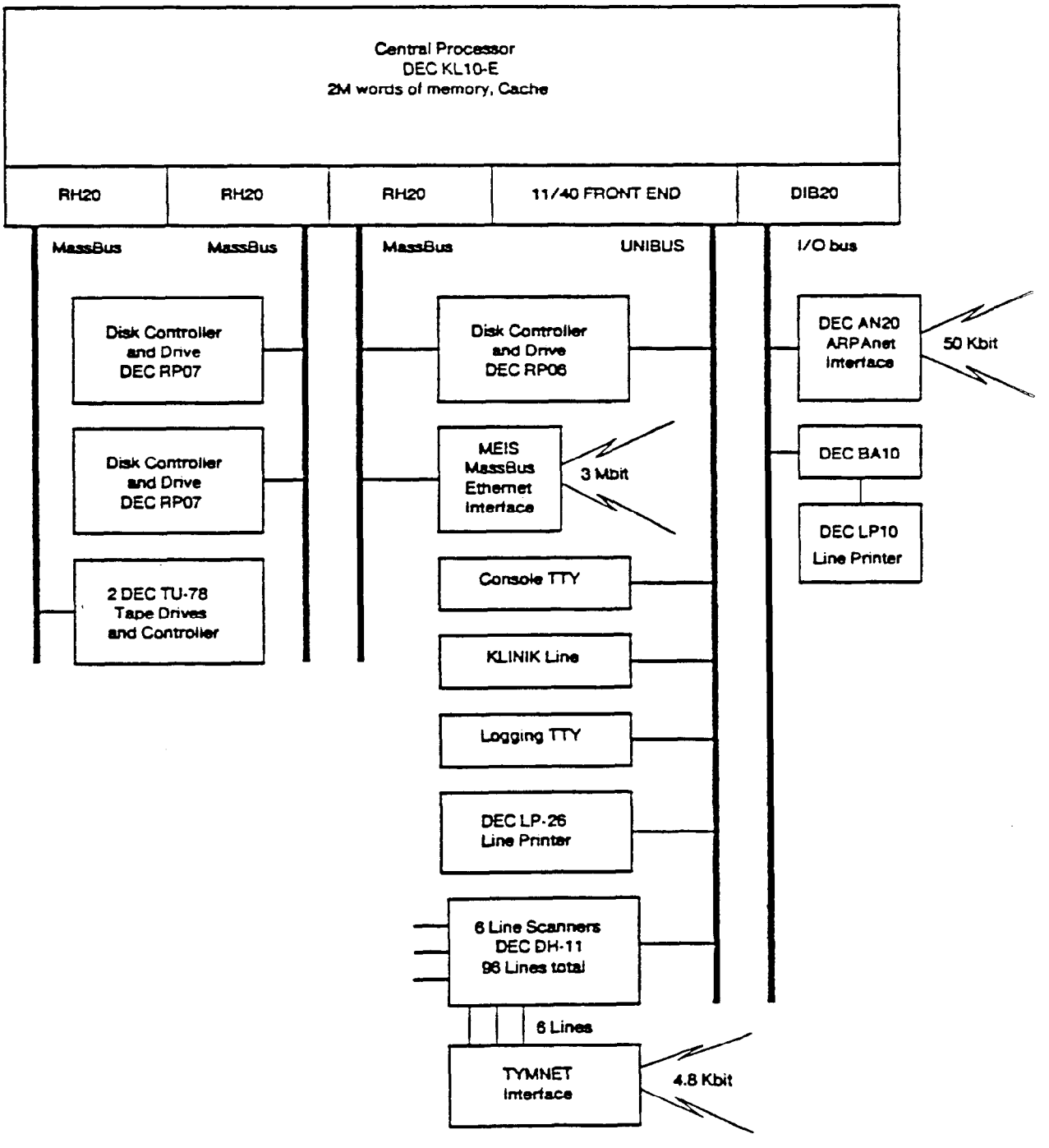

Figure 1: Current SUMEX-AIM Decsystem 2060 Computer Configuration

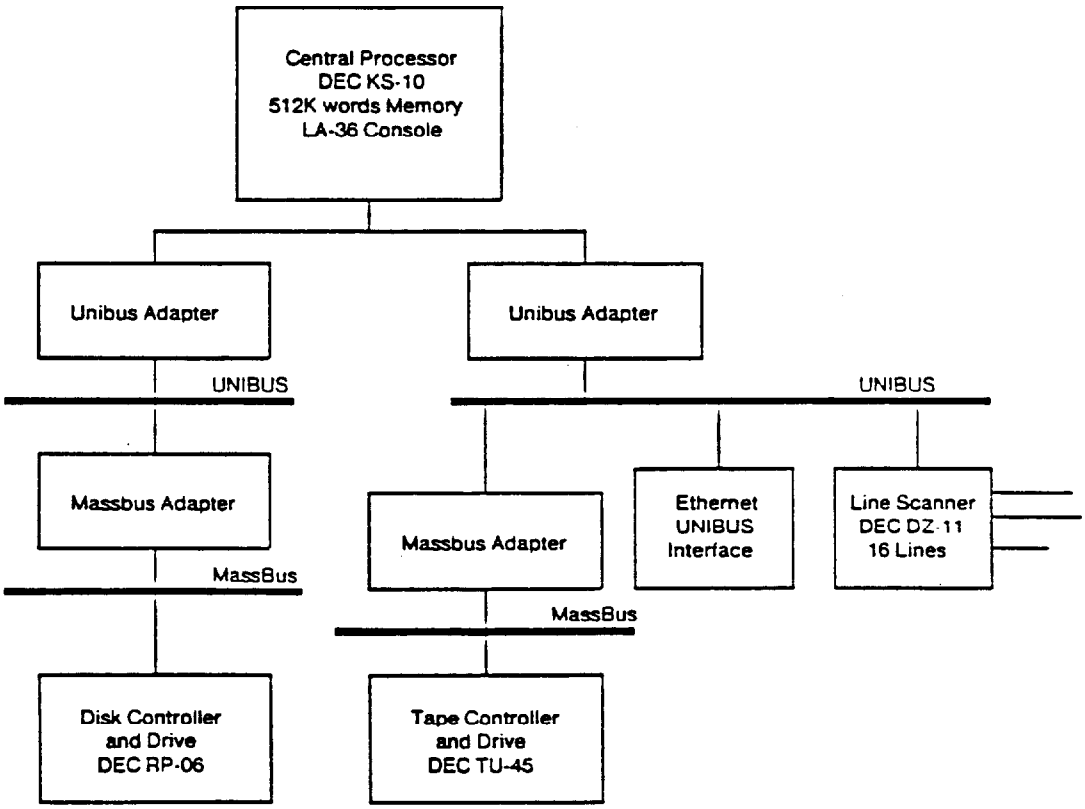

Figure 2: Current SUMEX-AIM 2020 Computer Configuration

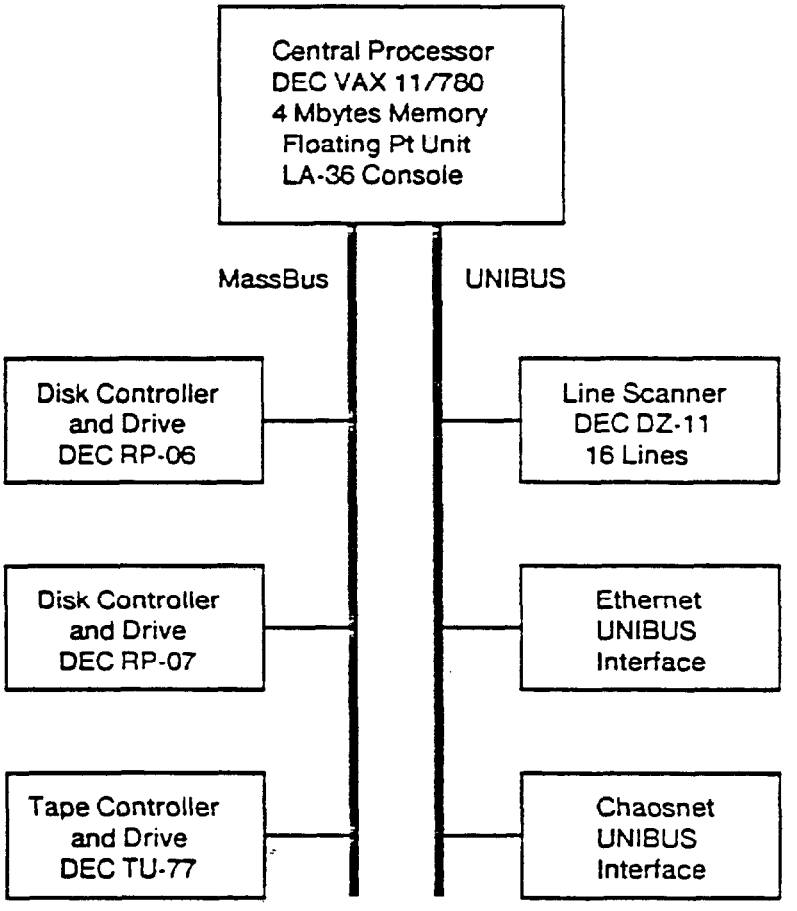

Figure 3: Current Shared VAX Computer Configuration

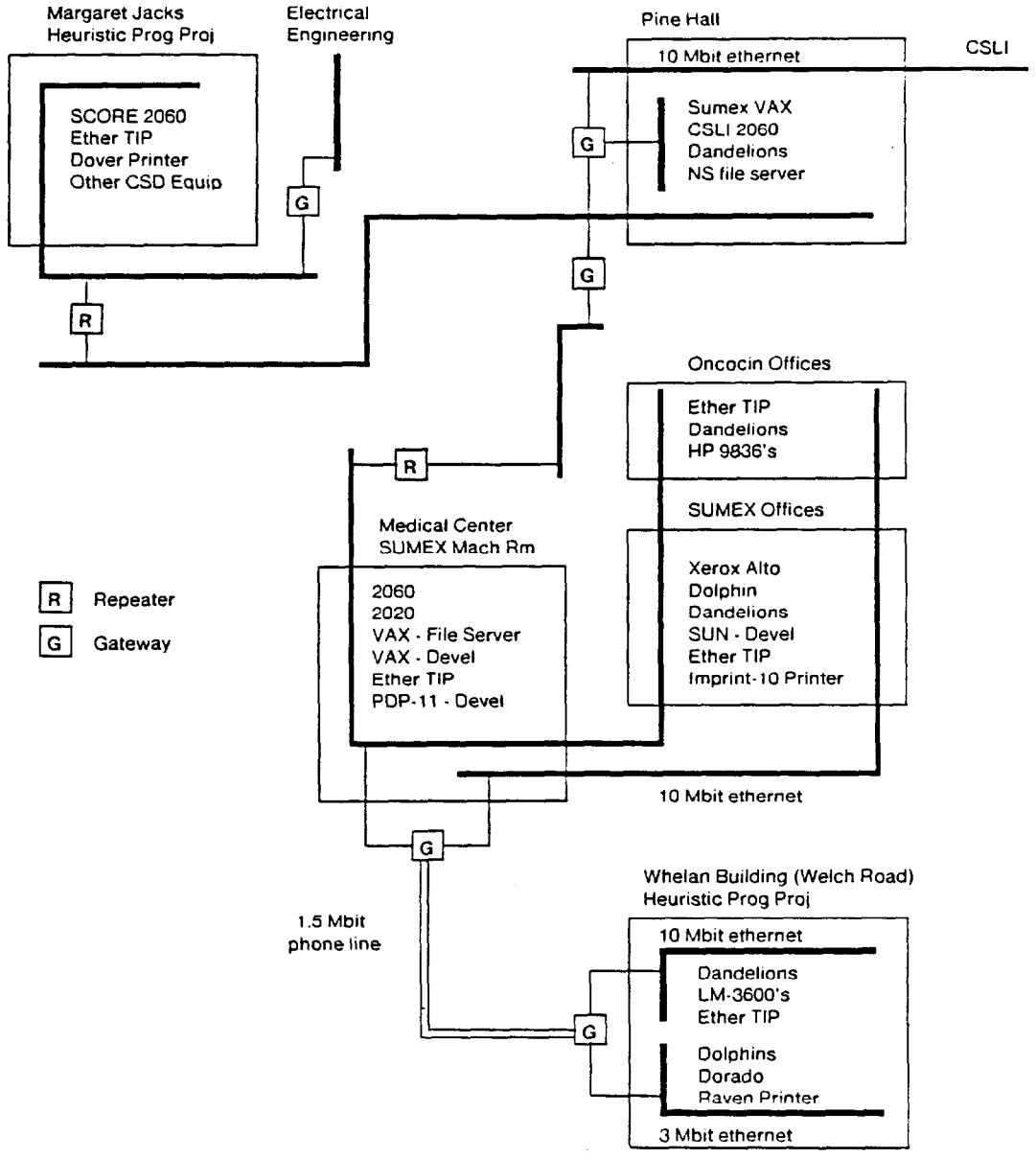

Figure 4: SUMEX-AIM Ethernet Configuration

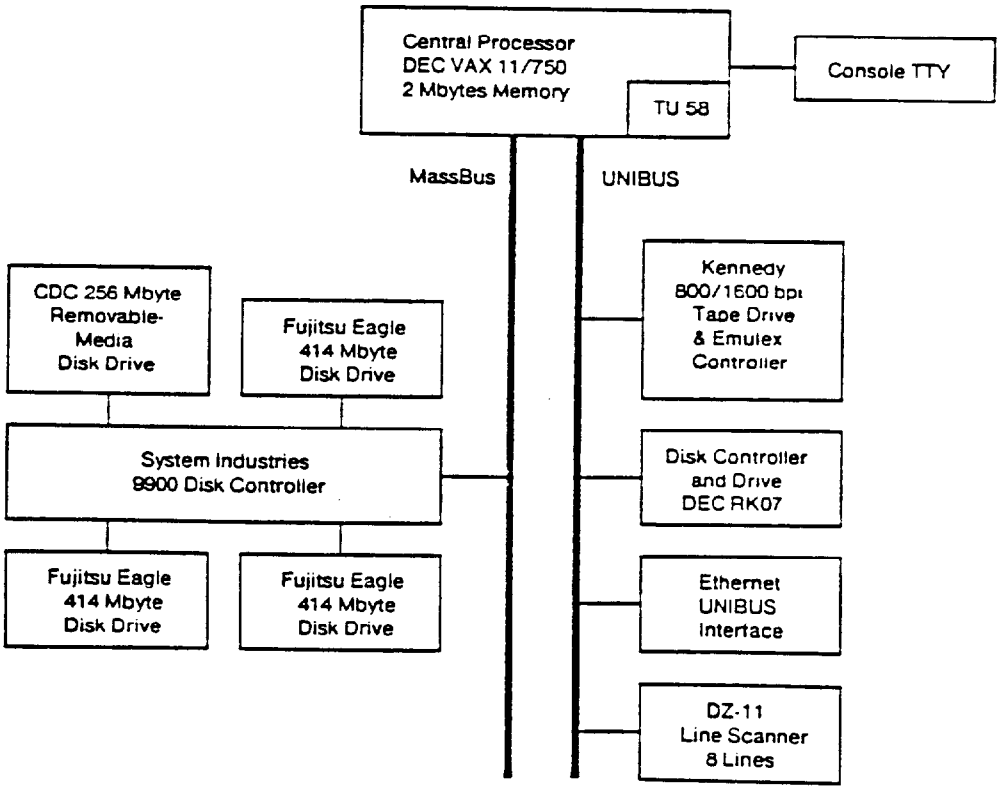

Figure 5: SUMEX-AIM File Server {SAFE}

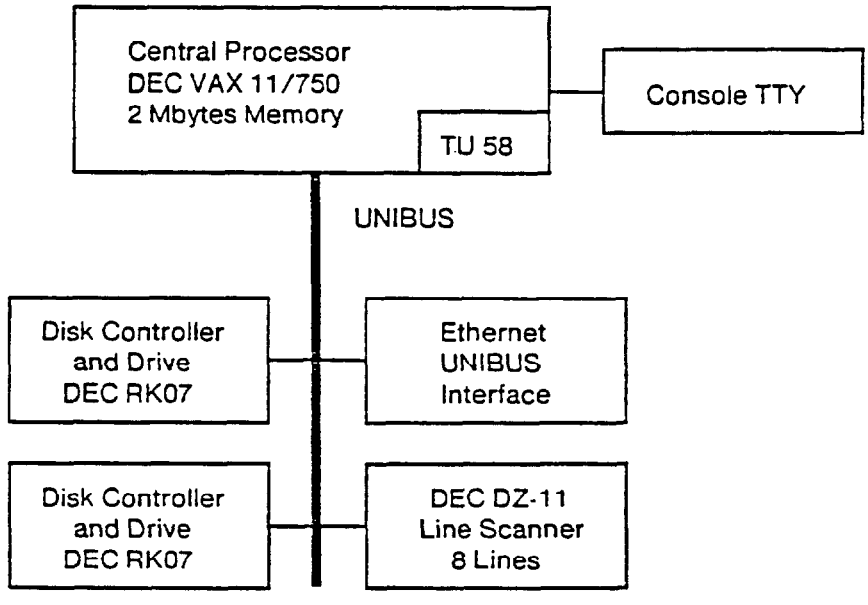

Figure 6: SUMEX-AIM Development Vax {ARDVAX}

I.A.2.3. Timesharing Systems

Continued support and development of our timesharing systems this past year has concentrated on several areas, including improvement of user services such as printing spoolers and archiving support, implementation of features from our KI10 Tenex system, enhancing network interface service, correcting encountered system bugs, and implementing new features for better user community support. In addition, we have invested further effort in supporting the VAX/UNIX system in conjunction with the SUMEX-AIM file server installation.

DECsystem 2060/TOPS-20 System

Support of our main timesharing machine, the DECsystem 2060, has continued during grant year 11.

- *Hardware development*

1. The DECsystem 2060 system now in operation at SUMEX differs greatly from our previous KI10 Tenex system. Whereas before, the KI10 system and TENEX software required much in-house development and support, life is easier with the 2060. Being at Stanford University, where there are at least 7 other DECsystem 20's with similar hardware and software is a great advantage. We are able to share our experiences with other sites, and have become an integral part of the Stanford DEC community. In addition, the DEC2060 hardware has been more reliable and easier to maintain than the KI10 system.

2. Additional modems were added to the 2060 to provide support for BELL 212A 1200 baud protocols. This adds an alternative to Vadic 1200 baud service. Modems which use the Bell 212A standard are more widely available for much less cost than Vadic modems.

- *TOPS-20 Monitor Software Enhancements*

1. A significant enhancement to our TOPS-20 monitor occurred this year when we implemented the software from our Tenex system which allowed extended support for the '?' feature of TOPS-20 when parsing filenames. This feature allows a user at any time to get a list of possible choices when needing to input a file name by just typing '?'. This returns an actual list of file names, whereas in the standard TOPS-20 monitor, just the string 'input filespec' was returned. This is a very significant and useful change to our TOPS-20 system.

2. We continued to keep up-to-date on the various bug fixes and monitor improvements that we received from DEC and other TOPS-20 sites. These included several fixes and rewrites to the Internet IP/TCP code which went under a major revision this past year.

3. We installed the capability for users to access their subdirectories as if they were the owners of such. While this may seem to be the logical way to implement subdirectories to begin with, DEC's models of subdirectories was a bit different. Our changes have since been installed on other DEC20's at Stanford and elsewhere.

4. We installed the capability to vary the allocation of windfall cycles in

accordance with user classes. This will allow us more flexibility in assigning jobs enough cycles to run comfortably while limiting their usage to a strict percentage of the machine.

5. We installed several new features in our TOPS-20 EXEC to facilitate the use of the system by the users. Among these features was the ability to edit any previous command you had entered, and then re-execute that command. This saves extra keystrokes and has proven to be very useful. The code to do this came from the University of Texas at Austin.

6. We switched our system this year to using encrypted passwords. This means that passwords are not stored in any readable form on the computer system, and if an illegal user should gain access to the system, he/she would not be able to find out the passwords of any other users. We feel this feature is quite important as the frequency of computer break-ins/attempts increases.

7. Software was added to our monitor in order to record the last reader of a file. Previously, only the ***date*** of the last read was recorded, while both the writer and date were recorded for creating and writing a file. This gives users the ability to determine which other persons may have been reading files.

- *Printer Support*

The support on the 2060 for various printers in the SUMEX community has been greatly enhanced this past year.

1. Support was added for the Xerox Raven printer at the Welch Road facility to allow spooling and direct output to the printer. In addition, code was added to the spooler to print out a header page identifying the user, filename, and date for each job.

2. Similar spooler support was added for the Xerox Dover printer in Margaret Jacks Hall.

3. SUMEX installed a Printronix line printer at Welch Road to allow users to print out files remotely from the 2060. The Printronix is connected to SUMEX via a twisted pair serial line.

4. We transformed TENEX software to the normal TOPS-20 line printer spooler program to look out for users who had accidentally printed an 'unprintable file', meaning a binary file of some sort which does not contain legible characters. We do this both by counting the number of binary characters in the first page of the file, and by not printing the file if the count exceeds a certain threshold. A similiar scheme is also used taking into account the vertical motion of the first page.

5. Additional modifications were made to the LP10 line printer driver software in the TOPS-20 monitor to improve the reliability of using this line printer, which came from our KI10 system.

6. We greatly enhanced our support of the IMAGEN Imprint-10 laser printer this past year. A new IP/TCP Ethernet interface was installed on the printer (discussed further in Section I.A.2.4 replacing the existing serial interface. This new interface allows for more efficient printer

operation, and greater flexibility in choosing output modes, number of copies, header pages, and other features. We have also implemented a TOPS-20 spooler for the Imagen as well.

- *User System Software*

1. We have continued to assemble and maintain a broad range of utilities and user support software on the 2060. These include operational aids, statistical packages, DEC-supplied programs, text editors, text search programs, file space management programs, graphics support, text-formatting and justification assistance, magnetic tape conversion aids, and many more. We also are importing software tools and packages wherever necessary to avoid reinventing the wheel and wasting our own efforts. Packages have been imported from Texas Instruments, Columbia University, the University of Texas at Austin, Yale University, and other Stanford sites.

2. SUMEX has continued to provide to its users the latest releases of various Lisp dialects that run under TOPS-20. This past year we agreed to provide disk space to store the 'official' version of Interlisp-20 from XEROX due to the fact that the machine at Xerox used to support Interlisp before was being removed. Interlisp-10 is now officially maintained at SUMEX by our staff and XEROX personnel. In addition, we continue to support the full variety of LISPUSERS packages. Portable Standard Lisp (PSL) developed at the University of Utah has also been installed on SUMEX.

3. We continue to use MM, a very powerful and flexible mail system, on the 2060. Electronic bulletin boards are also extensively-supported at SUMEX. These provide a rather informal mechanism for community discussions and debates. Other bulletin boards, read and contributed to throughout the INTERNET community, are available for perusal at SUMEX. These bulletin boards cover such topics as AI Discussions, Micro Computers, Terminals, and Workstations.

4. SUMEX participates with other Stanford sites in a general license for access to the SCRIBE text-formatting system from UNILOGIC, including versions to run under TENEX, TOPS-20, and UNIX. SCRIBE is the preferred tool for text preparation at SUMEX.

5. Versions of various user utilities and system utilities were updated throughout the year. These programs included network server processes, statistical packages, system daemon programs, and several programs for processing electronic mail.

6. Various system network tables and networking software were updated to accommodate the ARPANET split that occurred this year. The network change effectively split the ARPANET into two networks, the MILNET, which is a secure private part, and the rest of the ARPANET, which operates as the original ARPANET did before.

- *Documentation and Education*

We have expended considerable effort to develop, maintain, and facilitate access to our documentation so to accurately reflect available software. The

HELP and Bulletin Board subsystems 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 more difficult for users to track down all of the change pointers. Within manpower limits, we are in a continual process of reviewing the existing documentation system for compatibility with the programs now on line and to integrate changes into the main documents. This also will be done with a view toward developing better tools for maintaining up-to-date documentation.

- *Software Sharing*

1. As stated previously, 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.

2. The TENEX, TOPS-20, and UNIX 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 relationships between staff members of the various sites develop. These collegial interactions 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.

3. 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, experience with various hardware configurations, experience with new types of computers and operating systems, designs for local networks, 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 MM, SAIL, PASCAL, SOS, INTERLISP, the C compiler, VAX Ethernet code, the PHOTO program, ARPANET host tables, various user utilities, and many others.

4. Finally, we also have assisted groups that have interacted with SUMEX user projects in acquiring access to software available in our community. We are repeatedly providing tape preparation and copy service to many SUMEX-AIM projects to aid in sharing their software with outside requestors.

DECsystem 2020/TOPS-20 System

1. Monitor Upgrade -- Our 2020 system has continued to run very reliably this past year. We have updated the 2020 monitor with bug fixes and performance improvements regularly. There will likely be few further monitor releases for the 2020 since it does not support extended addressing and there are no plans to add this feature.

2. Demo Controls -- We continue to use the 2020 system for demos of AI systems developed at SUMEX. This demo system takes advantage of the "pie-slice" scheduler in the TOPS-20 release 4 monitor. We now guarantee dedicated users a large fraction of the machine but also allow others to do useful work when the demo demand is low. This system has nicely met the needs of both groups.

VAX/UNIX Systems

We continued to provide systems support for the VAX/UNIX 11/780 system (named 'AIMVAX') shared by the SUMEX-AIM community and Stanford Computer Science Department. Various efforts included supporting the UNIX monitor, installing new network software, and in bringing up various user subsystems.

Further development has continued in support of the SUMEX-AIM File Server (SAFE) based on a VAX 11/750. We successfully converted SAFE to Berkeley Unix 4.2 server, and with the help of the Computer Science Department, converted the Ethernet Pup software to run under UNIX 4.2

I.A.2.4. Professional Workstations

Our ongoing movement to professional workstations is taking on several forms. We continue to carry out our acquisition plans for acquiring Lisp machines for use by the SUMEX community, as well as investigating the use of remote virtual graphics and new lower cost workstations in our environment. This work is prototypical of what other groups will face and we hope will serve to find effective solutions to common problems.

Lisp-based Scientific Workstations

SUMEX carefully developed and implemented our equipment acquisition plan for year eleven by buying seven Xerox 1108 Lisp machines for use by SUMEX-AIM projects. Two of these machines were purchased with special upgrade packages to provide floating point capability, expanded microcode, and expanded memory. Our experiences with these machines will be reported in next years report.

The XEROX Dolphin on loan to Rutgers University was returned to SUMEX this year. This Dolphin had effectively served the Rutgers-AIM community in setting up their Ethernet network and provided initial exposure to the Lisp machine technology. Now that that experiment is successfully completed, the Dolphins will be used for AI system development at Stanford.

SUMEX installed two SYMBOLICS 3600 Lisp machines, purchased with DARPA funding, for use within the Heuristic Programming Project (HPP) at their new location at Welch Road. We are currently awaiting a new release of the Symbolics Operating System software before we can provide Ethernet access to our file server from these machines. The 3600's are used regularly by members of the HPP.

We still are using 4 preproduction models of the Dolphin workstations. One preproduction model has been exchanged for a production system, and we are on schedule with XEROX to exchange the remaining 4 machines for production models at no extra cost. This process is hampered by the rate at which XEROX themselves can get production machines.

We studied the benefits of buying the extended memory and microstore upgrades to the Xerox 1108 Dandelion announced at AAAI-83 as being "under development." We concluded that some users would benefit greatly from these enhancements and others not at all. The most marked improvements came from system which were extremely memory limited, such as NEOMYCIN. SUMEX will be acquiring two 1108's with the upgrades for full time use and testing.

A close relationship between SUMEX and the newly-formed Center for the Study of Language and Information (CSLI) at Stanford was established. This has already benefited SUMEX (and the ONCOCIN project in particular) in the loan, by CSLI to SUMEX, of two Xerox 1108's which have been in constant use by researchers since January 12th. The SUMEX staff assisted CSLI in bringing their DecSystem20 and network environment on-line. CSLI has informally expressed an interest in working on the problem of distributed AI computation with SUMEX researchers. CSLI will have 110 1108's on the Ethernet within the year. This resource suggests some exciting solutions to former compute-bound problems. The ONCOCIN group has already implemented a preliminary network-based Interactor which permits elements of ONCOCIN to run concurrently on different machines. As of this writing, the Reasoner and its Debugger have been made to run transparently in this mode, and to make good use of both processors.

Virtual Graphics

SUMEX continued the development of a Virtual Graphics system written in Interlisp-10 and running on our 2060. Any user running the V system on a workstation can then use the package on the 2060 to drive the graphics display on the workstation. A current application is to take nuclear magnetic resonance data on the 2060 and display the atoms and their bonds on a SUN workstation by using splines. This development is in its infancy, but is opened ended and has great potential with the price of workstations capable of decent graphics reaching the two to three thousand dollar range. It allows those users who cannot afford expensive lisp machines to have full graphics power available to them by doing the actual graphics applications on a large time shared system, and then doing the graphics itself remotely on a less expensive workstation. This development can help users take advantage of the computing power of the DECsystem 20, while providing many of the high speed graphics advantages of the Lisp Machines.

Apple Workstation Development

SUMEX-AIM has initiated a development project to pursue the effectiveness and possible use of low cost personal workstations within our environment. After examing a number of new personal computers and workstations on the market, such as the Hewlett Packard 150, IBM PC, Sun workstations, and others, we chose the Apple Macintosh and Apple Lisa on which to begin our work in this area. These machines were chosen technically due to their built in graphics, networking, mouse, windows, and menu support. We also considered the very beneficial relationship formed between Apple and Stanford University which provides us direct access to Macintosh hardware and software documentation which is a necessity for the type of work we plan to do.

Our Macintosh development encompasses several areas ;

1. *INFO-MAC Discussion List*

   An electronic discussion list was originated at SUMEX, and is currently maintained here, to foster sharing and communication among research groups and universities that are interested is pursuing the serious use of the Macintosh within their respective environments. This list has been highly successful in collaborating on Macintosh development and the sharing of ideas. The discussion list currently contains over 50 sites, and well over 1000 participants.

2. *C Development Environment*

   A vital link in our development of Macintosh software is creating a C based development system on our VAX computers for the coding and downloading of software. Utilizing existing MC68000 C cross compilers on our VAX, we are developing the necessary linkages in order to make the appropriate system calls to routines in the Macintosh ROM's for sophisticated graphics and system related functions.

3. *Macintosh print servers*

   In order to effectively use the Macintosh as a stand-alone workstation, we will provide the ability to print out Macintosh developed files on our IMAGEN laser printers.

4. *Applebus to Ethernet Interface*

   This development involves the hardware and software necessary to be able to

access various file and print servers on our Ethernet from a Macintosh. The Macintosh will be connected to the Apple network called Applebus. Our hardware provides an interface between Applebus and 10MB Ethernet. The software necessary for this project involves formulating Macintosh file level and block level I/O requests into properly formatted Internet packets.

5. *Virtual Graphics on a Lisa*

   The Virtual Graphics System, as previously reported, is in great need of a low cost workstation on which it can run. We have started a project to port the Virtual Graphics system to a Apple Lisa in hopes of providing to our users high speed graphics at remote locations. We will report further on this project in next year's report.

Anticipating the popularity of our Macintosh developments, we are fully prepared to make our efforts available to other research sites, Universities, and non-profit institutions on a royalty-free basis in hopes of fostering continued development and communal sharing.

In addition to Lisp-based scientific workstations, we believe the use of low cost workstations, which offer suitable local processing power, high resolution screens with easy to use user interfaces, and networking and communications abilities, are vital to the future of our resource. Our Macintosh and Lisa development efforts will allow us to use and experiment with these workstations in our environment.

Hewlett Packard Development

SUMEX assists the ONCOCIN Project is developing a computing environment for developing AI applications based on HP 9836 workstations. These workstations were part of a gift from Hewlett Packard to the Oncocin Project. Additional support peripherals for the 9836's included large capacity disk drives, color monitors, graphic tablets, and a laser printer. Work is proceeding to network these machines onto the SUMEX Ethernet as soon as suitablke networking hardware is available from HP. These machines will be used for new and existing projects within Oncocin.

I.A.2.5. Networking and Communications

A highly-important aspect of SUMEX-AIM is effective communication with remote users and between the growing number of machines available within the SUMEX resource. In addition to the economic arguments for terminal access, networking offers other advantages for shared computing, including improved inter-user communications and more effective software sharing.

Users accessing a remote computer will use a hardline connection to the computer as a standard of comparison. Local networks stand up well in this comparison but remote network facilities do not. Data loss is not a problem in most 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, remote networking relies upon shared use of communication lines for 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. We continually monitor the load statistics for our direct, dialup, and TYMNET lines to avert logjam situations.

TYMNET

TYMNET provides broad geographic coverage for terminal access to SUMEX from throughout the country and increasingly from foreign countries. With the installation of our new DEC2060 computer system in January of 1983, we installed new TYMNET equipment. After the initial debugging of the new equipment (called TYMCOM) the equipment has been quite reliable. However, some months after the installation it was discovered that the XON/XOFF protocol between the Tymcom and the 2060 had not been properly specified in the Tymcom and was corrected. The number of user complaints about connection problems have been greatly reduced. This is thought to be the result of improved "backbone" lines within Tymnet and the installation of triple-duty modems which simplify things for the users.

ARPANET

We retain our advantageous connection to the Department of Defense's ARPANET, now managed by the Defense Communications Agency (DCA). This connection has facilitated close collaboration with the Rutgers-AIM facility and many other computer science groups that are also on the net. 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. The ARPANET is also key to maintaining on-going intellectual contacts between SUMEX projects such as the Stanford Heuristic Programming Project authorized to use the net and other active AI research groups in the ARPANET community.

This past year, SUMEX-AIM participated in the split of the ARPANET into two networks; the MILNET, which is a highly secure strictly DOD-related part of the network, and the ARPANET, which is the remainder of the ARPANET sites. This latter net functions as we knew the ARPANET before. The MILNET can only be accessed via mail gateways. No TELNET or FTP to MILNET sites is allowed. In addition, access to the ARPANET TAC's (Terminal Access Controllers) was restricted this past year to only those users who were granted TAC access cards, which meant their username was registered with the Network Information Center, and they were given a password with

which they could dial into the ARPANET. SUMEX arranged for guest cards for those users who needed such access.

We continue to be called upon to interact with outside organizations which are (or wish to be) connected to our IMP. The line to Advanced Information and Decision Systems occasionally causes trouble requiring diagnosis. The intended connection to Perceptronics Inc. has evidently been canceled.

ETHERNET

A substantial portion of our system effort this past year went into continued development of local Ethernet facilities which link the SUMEX resource hardware with other parts of the campus, namely to 701 Welch Road, which is the new location of the Heuristic Programming Project, and to the Computer Science Department building on campus. We have also invested a great amount of effort this year to begin our transfer to a 10 megabit Ethernet, while continuing support of our current 3 MB ethernet.

Specific areas of Ethernet development include:

1. *Leaf server* -- We continued support of the Sequin reliable packet protocol and Leaf byte-level file transfer protocol to enable our Xerox D machines to access files on our DEC20 systems. The Leaf server had to be modified on the DEC20 this past year when we switched to using encrypted passwords. The LEAF server implementation for the 4.2 BSD release of Unix was also debugged and installed at SUMEX. This allows us to access files stored on our VAX file servers from either our 10MB or 3MB networks.

   The Leaf protocol is built into the lowest levels of the Dolphin I/O system, and allows any file on a remote file server to be accessed as easily as a disk file in both paged or random access mode. The latest updates to the Sequin transport level have made marked improvements in efficiency. The 2020 now performs Leaf file transfers with a speed approaching that of XEROX's dedicated file server.

2. *TOPS-20 Ethernet Server* -- We continued to maintain and improve the Ethernet service under the TOPS-20 operating system. This included updates to the TELNET and FTP programs, as well as mail software, the previously mentioned Leaf server, and network table maintanence programs.

3. *Ethernet Gateway* -- Our Ethernet gateway software has continued to run reliably and effectively. The previous problems with lost packets and delayed terminal response has been fixed, the cause of which was a bad memory board and a software bug in the TOPS-20 operating system. Serious problems that affected our net connectivity to other parts of campus were also discovered and repaired this past year, thereby providing us with over 99% net connectivity to the rest of campus. The changes involved board repair and modifications to the topology of the campus Ethernet.

   The gateway itself was generalized to handle 3 or more directly connected networks where previously it had only dealt with 2 such networks. We currently have two gateways, each handling the traffic between three local networks, two of which are 3MB and one a 10MB network.

4. *Ethernet TIP (EtherTIP)* -- The EtherTIP provides multiple terminal access to the Ethernet. A PUP ethernet operating system was written for MC68000-based processors, and a MC68000-based EtherTIP was built based on this.

The EtherTIP software has undergone further enhancements in the past year. Portions of this work was done in conjunction with the Stanford Computer Science Department. Among those enhancements are the following:

a. It now accepts incoming connections to line printer ports, and for remote system diagnosis.

b. It can simulate the "old" Stanford EtherTIP for users who have not yet made the transition to the new environment.

c. The user interface is more flexible to suit the needs of an increasingly diverse user community.

The EtherTIP software has developed into a very stable system, and one enjoying good use within the SUMEX community.

5. *10 MB/SEC Ethernet Development* -- SUMEX made a major move this past year to begin our transfer to a 10 megabit/sec network. While the current 3 megabit/sec network continues to serve us well, many new workstations and printers are coming on the market with only 10 MB/SEC interfaces, and in addition, since 3 MB/SEC networks were only used a very few selected settings, it is becoming increasingly difficult to find replacement parts when failures do occur.

   Therefore, this past year saw several efforts involved in installing and supporting the SUMEX 10MB/SEC Ethernet ;

   a. Reworking the entire Ethernet system software to handle ***both*** 3 and 10 megabit link level standards, i.e., addressing and encapsulation are transparent to the user levels. We similarly made the network link level protocols transparent to the the user level software. In this way one can communicated using PUP protocols on a 10MB/SEC ethernet and the user software does not have to change.

   b. Adding address resolution protocols for PUP and IP so that the 3MB/SEC byte addresses can be translated to 10 MB/SEC hardware addresses for the link level. This enables one to communicate using PUP or IP between 3 and 10 megabit hosts.

   c. Integrating XNS and IP into the PUP routing mechanism.

   d. Solving some rather subtle software/hardware integration problems in order to simulate "ethernet" on the HPP/Welch Road "twisted pair" ethernet.

   e. Bringing up the 3 MB/SEC EtherTIP on the 10 MB/SEC network was a proof that the above worked. It was done without any changes to the TIP software itself by simply relinking it with the 10 MB/SEC system software. This required only one additional piece of logic. When a 10 MB/SEC host wants to communicate using PUP which is a 3 MB/SEC protocol, then it must find its PUP address from some host on the 10 MB/SEC network. The gateway maintains a translation table, and listens for such requests, thus translating the 10 MB/SEC hardware address into a "soft PUP address," and replying to the requesting host.

6. *HPP-SUMEX Communication link* -- The Heuristic Programing Project

(HPP) relocated from its campus location to 701 Welch Road, adjacent to the Stanford campus. Since this group is a primary user of the SUMEX computer facility and the principal focus for core AI research, a communication link between the new location and SUMEX machine room was imperative. Several communication schemes for establishing a reliable and relatively fast link were considered, namely ; microwave, laser, infrared, direct ethernet (by trenching and placing a direct ethernet cable), ATT's T1 service and others.

All of the above schemes would have necessitated large budgetary outlays and some would have imposed lengthy time delays (getting permits and the like) due to jurisdictional boundaries. The idea of using bare copper telephone pair already in place looked very attractive especially if reasonable speed and reliability could be achieved. The wire distance between the above mentioned locations is approximately 2000 ft. A design goal was established to try to develop a communication link with Ethernet type speed (3MB/SEC) between these two locations.

Utilizing high driving capacity drivers (differential) and ultra high speed, high sensitivity receivers a transceiver was designed and tested for maximum transmission speed with maximum reliability. The final configuration resulted in a half duplex transmission over a bare copper twisted pair in each direction utilizing Manchester coding at a reliable transmission speed of 1.25MBs/sec. each direction for an aggregate speed of 2.5MBs/sec. This communication link has been in operation for about six months now without any appreciable down time or noticeable error rate or data delays. Many HPP researchers are utilizing this link to communicate with SUMEX and the University Ethernet network. In addition, various Lisp machines and printers located in the HPP facility and connected to a local network there are also able to communicate with the University network.

INTERNET SOFTWARE

One major issue we face at SUMEX-AIM in support of our network environment is the lack of standardization in network protocols among various vendors. Currently, many vendors are adding support to their products for the Internet (IP/TCP) protocols. SUMEX continues to support the IP/TCP protocols on the DEC2060, and we are currently alpha-testing a release of Interlisp-D which also supports IP/TCP protocols. In addition, we sucessfully adapted the IP/TCP software to our VAX systems running UNIX 4.2BSD. This Vax TCP adaptation involved provisions for subnet routing, 3 MB/SEC byte swap problems, encapsulation problems and 10 MB/SEC debugging with our gateways.

I.A.2.6. Progress in Core Research

Over the past year we have continued to support several core research activities aimed at developing information resources, basic AI research, and tools of general interest to the SUMEX-AIM community. SUMEX is providing only partial support for these projects, with complementary funding coming from ARPA, ONR, NLM and NSF contracts and grants to the Stanford Heuristic Programming Project.

Core Research

Core Research at SUMEX-AIM focuses on understanding the roles of knowledge in symbolic problem solving systems, its representation in software and hardware, its use for inference, and its acquisition. We are continuing to develop new tools for system builders and to improve old ones. The research crosses a number of application domains, as reflected in the subprojects discussed earler, but the main issues that we are addressing in this research are those fundamental to all aspects of AI. We believe this core research is broadening and deepening the groundwork for the design and construction of even more capable and effective biomedical systems.

As mentioned above, although our style of research is largely empirical, the questions we are addressing are fundamental. The three major research issues in AI have, since its beginning, been knowledge representation, control of inference (search), and learning. Within these topics, we will be asking the following kinds of questions. As our work progresses, we hope to leave behind several prototype systems that can be developed by others in the medical community.

1. Knowledge Representation -- How can we represent causal models and structural information? What are the relative benefits of logic-based, rule-based, and frame-based systems? How can we represent temporal relations and events so that reasoning over time is efficient?
2. Knowledge Acquisition -- How can an expert system acquire new knowledge without consuming substantial time from experts? Can we improve the knowledge engineering paradigm enough to make a difference? Can automatic learning programs be designed that will work across many disciplines? Will cooperative man-machine systems be able to open the communication channel between expert and expert system?
3. Knowledge Utilization -- By what inference methods can a variety of sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? What is the nature of strategy and control information?

Plans for the Coming Year

Several systems have been developed in recent years to serve as vehicles for knowledge engineering and research on knowledge representation and its use. Knowledge acquisition (including machine learning) and advanced architectures for AI will be the two areas of most new activity in the coming year. Research on these topics obviously must draw on on-going work in representation and control.

In particular, we will focus on

- Inductive learning of MYCIN-like rules from case data in the domain of diagnosing disorders where the chief complaint is jaundice;
- Learning from experience in domains where the means for interpreting new

data are largely contained in the emerging (and thus incomplete and not wholly correct) theory;

- Learning by watching a medical expert diagnose cases presented by NEOMYCIN;
- Investigating complex signal understanding systems for ways to exploit and represent concurrency with a view toward hardware and software architectures that may be capable of several orders of magnitude improvement in performance.

Further information on the core research at SUMEX-AIM and the Heuristic Programming Project can be found in the Projects section starting on page 89.

I.A.2.7. Resource Operations Statistics

The following data give an overview of various aspects of SUMEX-AIM resource usage. There are 5 subsections containing data respectively for:

1. Overall resource loading data (page 31).
2. Relative system loading by community (page 33).
3. Individual project and community usage (page 36).
4. Network usage data (page 44).
5. System reliability data (page 44).

For the most part, the data used for these plots covers the entire span of the SUMEX-AIM project. This includes data from both the TENEX KI10 system and the current DECsystem 2060. At the point where the SUMEX-AIM community switched over to the 2060 (February, 1983), you will notice severe changes in most of the graphs. This is due to many reasons which I will mentioned briefly here ;

1. Even though the Tenex operating system used on the KI10 was a forerunner of the current Tops20 operating system, the Tops20 system is still different from Tenex is many ways. Tops20 uses a radically different job scheduling mechanism, different methods for computing monitor statistics, different I/O routines, etc. In general, it can not be assumed that statistics measured on the Tenex system correlate one to one with similar statistics under Tops20.
2. The KL10 processor on the 2060 is a faster processor than the KI10 processor used previously. Hence, a job running on the KL10 will use less CPU time than the same job running on the KI10. This aspect is further complicated by the fact that the SUMEX KI10 system was a dual processor system.
3. The SUMEX-AIM Community was changing during the time of the transfer to the 2060. The usage of the GENET community on SUMEX had just been phased out. This part of the community accounted for much of the CPU time used by the AIM community. Since the purchase of the 2060 was partially funded by the Heuristic Programming Project (HPP), an additional number of HPP Core Research Projects started using the 2060, increasing the Stanford communities usage of the machine. And finally, the move to the 2060 occurred during a pivotal time in the community when more and more projects were either moving to their own local timesharing machines, or onto specialized Lisp workstations. It also was the time for the closure of many long time SUMEX-AIM projects, like Dendral and Puff/VM.

Any conclusions reached by comparing the data before and after February, 1983 should be done with caution. The data is included in this years annual report mostly for casual comparison. Starting next year, only data from the 2060 will be recorded on the annual report. Readers will be referred to previous annual reports (such as this one) for data from the KI10 Tenex system.

Overall Resource Loading Data

The following plots display several different aspects of system loading over the life of the project. This data includes usage of the Tenex KI10 system and the current DECsystem 2060.

These plots include total CPU time delivered per month, the peak number of jobs logged in, and the peak load average. The monthly "peak" value of a given variable is the average of the daily peak values for that variable during the month. Thus, these "peak" values are representative of ***average*** monthly loading maxima and do not reflect the largest excursions seen on individual days, which are much higher.

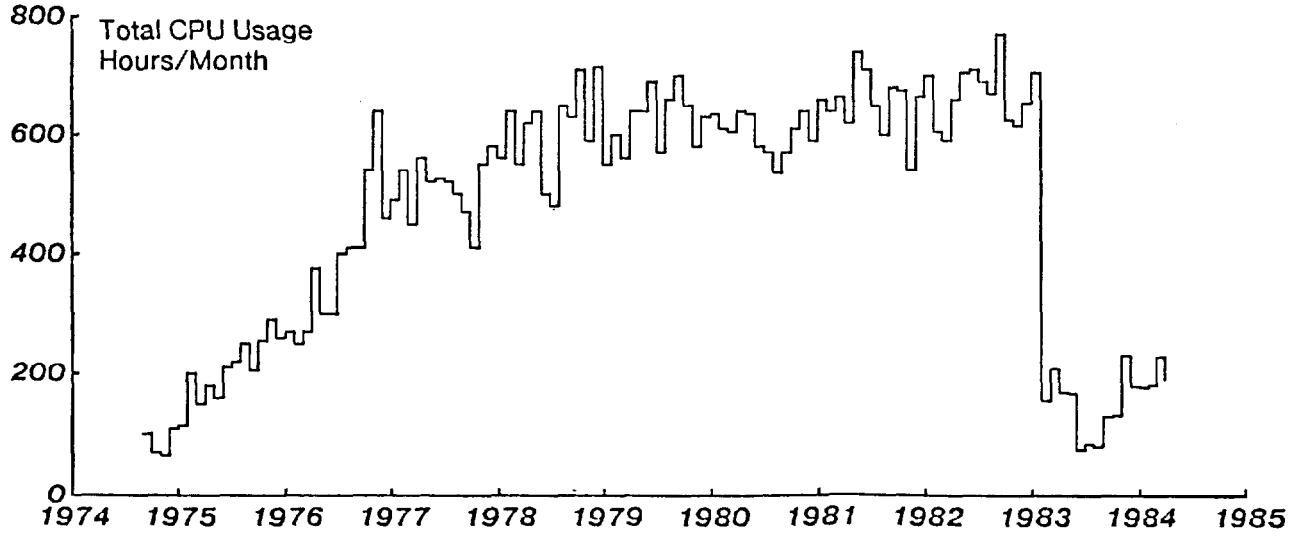

Figure 7: Total CPU Time Consumed by Month

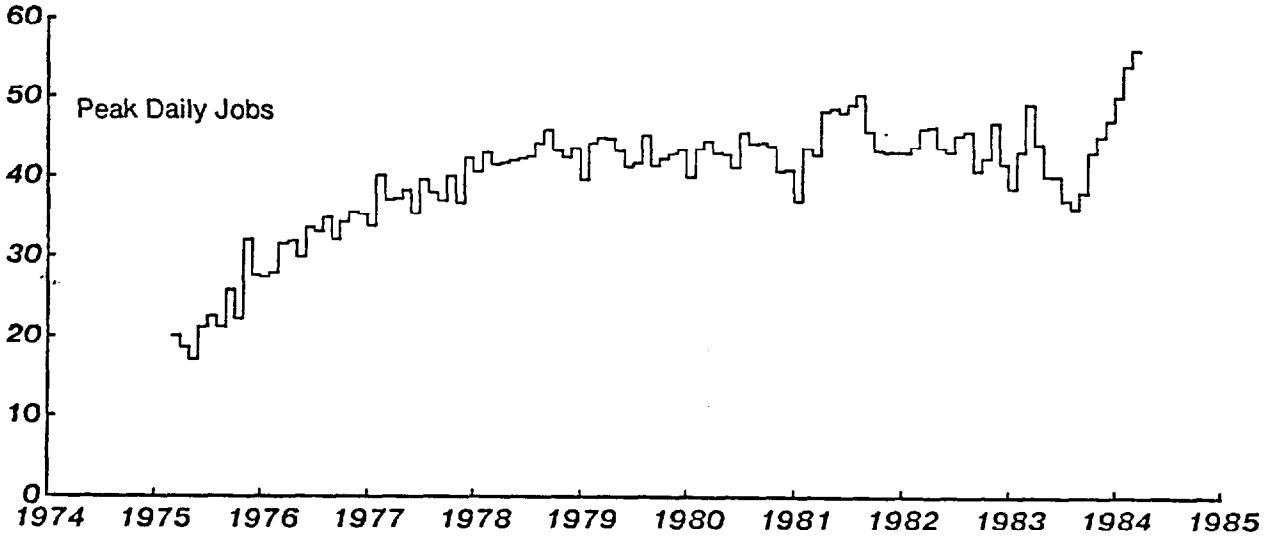

Figure 8: Peak Number of Jobs by Month

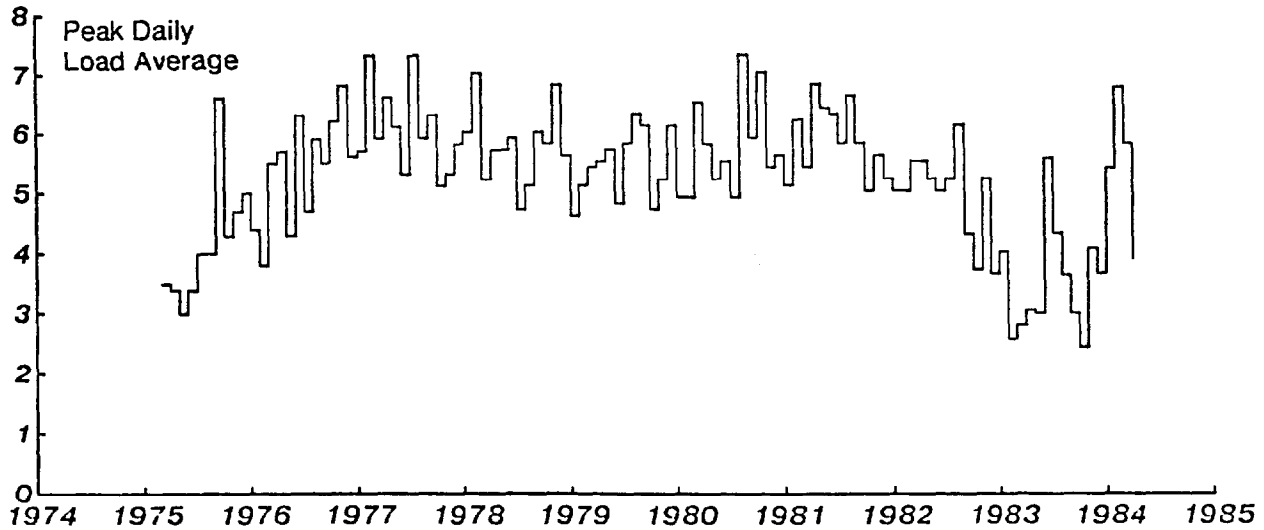

Figure 9: Peak Load Average by Month

Relative System Loading by Community

The SUMEX resource is divided, for administrative purposes, into three major communities: user projects based at the Stanford Medical School (***Stanford Projects***), user projects based outside of Stanford (***National AIM Projects***), and common system development efforts (***System Staff***). As defined in the resource management plan approved by the 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 resources and terminal connect time for each of these three communities relative to their respective aliquots is shown in the plots in Figure 10 and Figure 11. As mentioned on page 30, these plots include both KI10 and 2060 usage data.

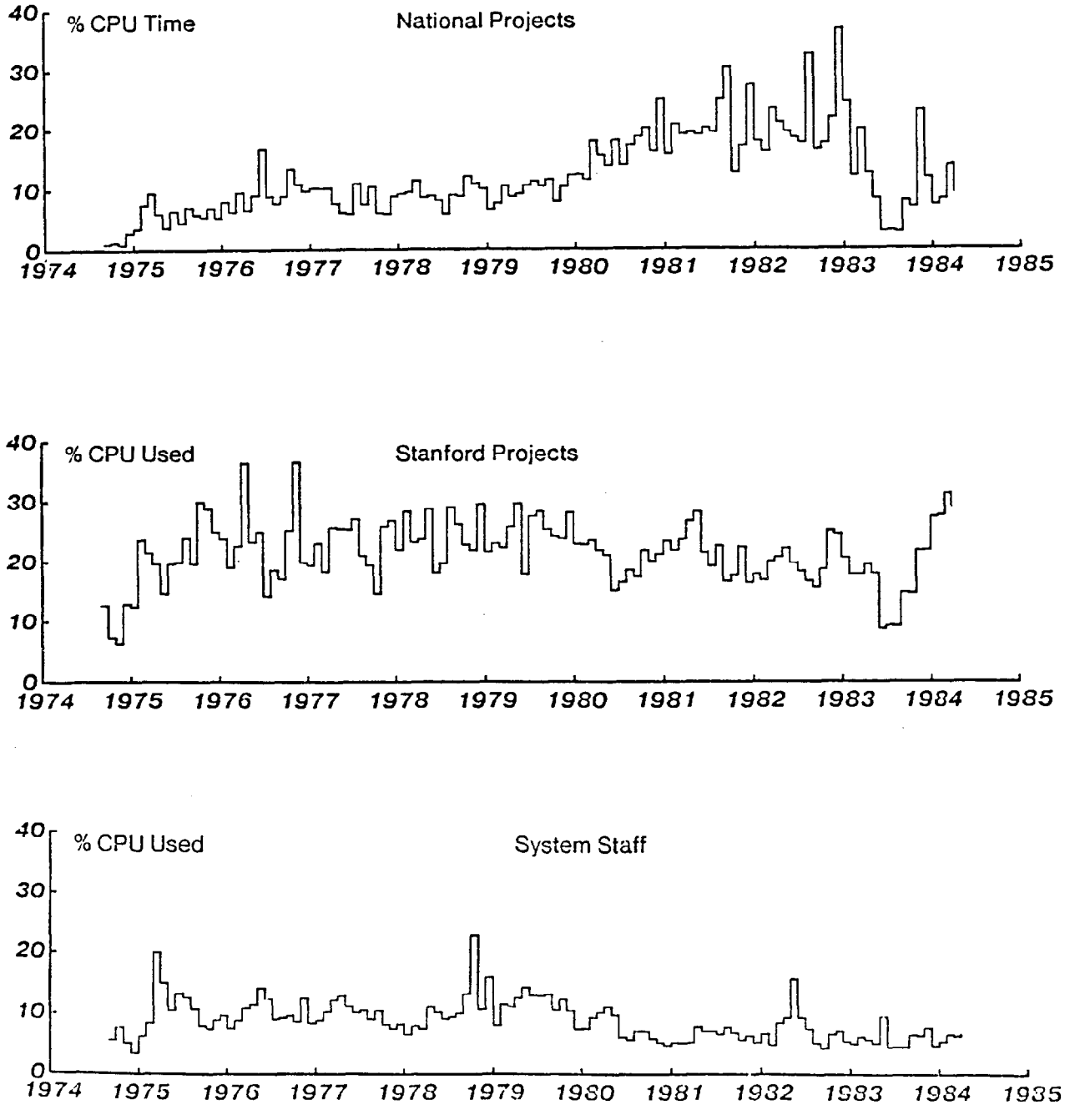

Figure 10: Monthly CPU Usage by Community

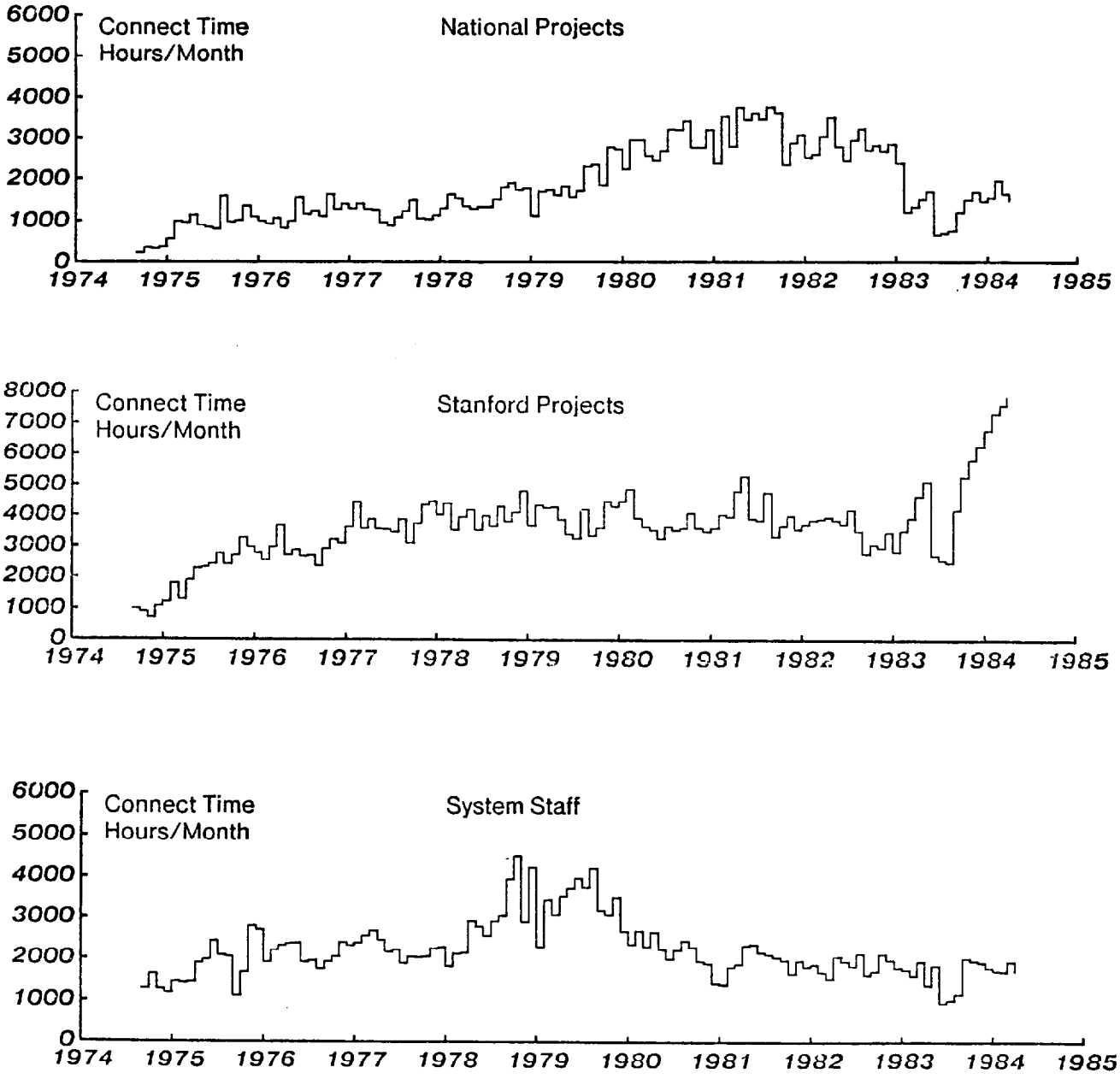

Figure 11: Monthly Terminal Connect Time by Community

Individual Project and Community Usage

The following histogram and table show cumulative resource usage by collaborative project and community during the past grant year. The histogram displays the project distribution of the total CPU time consumed between May 1, 1983 and April 30, 1984, on the SUMEX-AIM DECsystem2060 system.

In the table following, entries include a text summary of the 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, 1983 and May, 1984.

Several of the projects admitted to the National AIM community use the Rutgers-AIM resource as their home base. We do not explicitly list these projects in this annual report covering the Stanford SUMEX-AIM resource. We do record information about the Rutgers resource itself, however, and note its separate resource status with the flag "[Rutgers-AIM]".

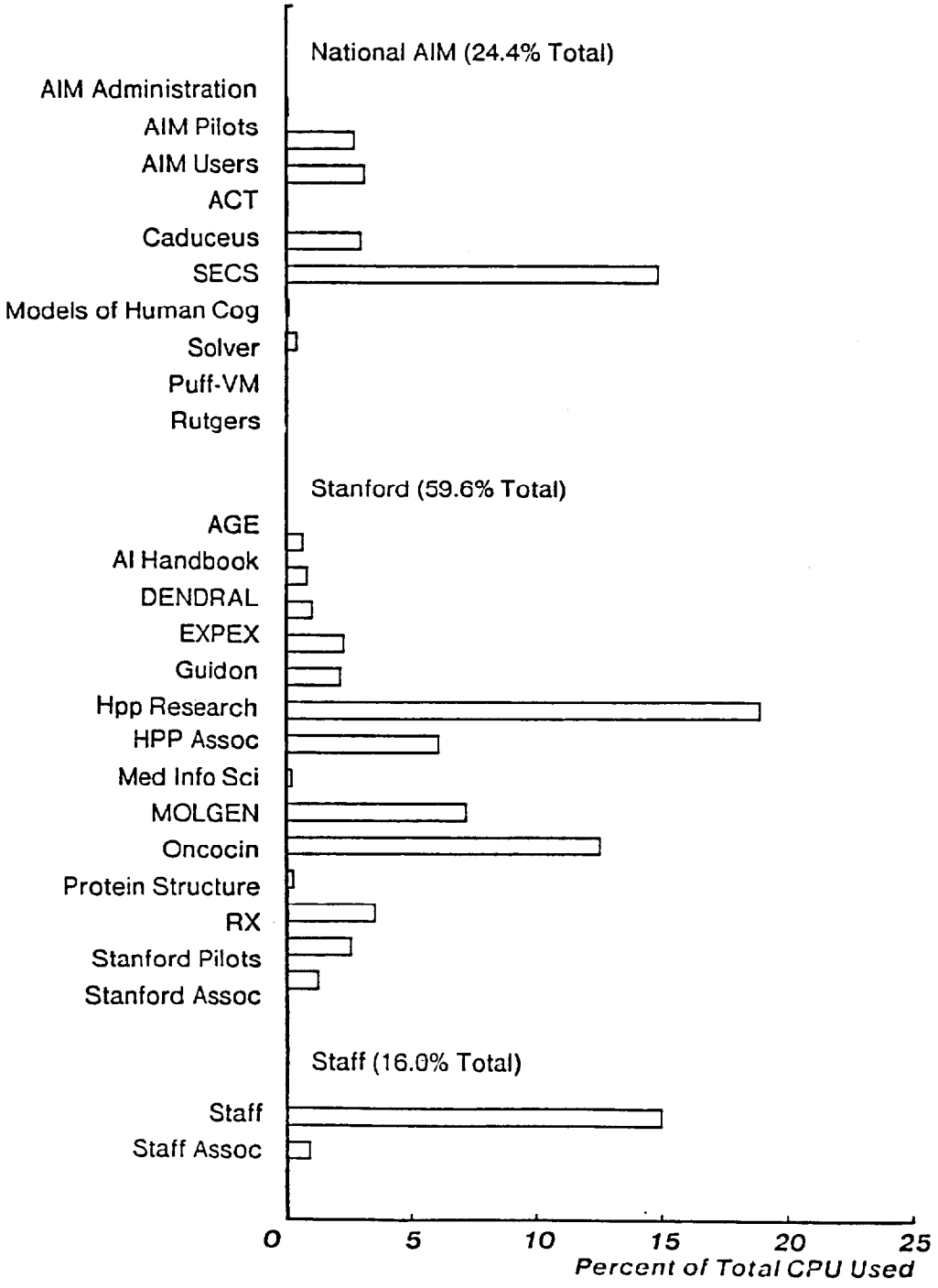

Figure 12: Cumulative CPU Usage Histogram by Project and Community

Resource Use by Individual Project - 5/83 through 4/84

National AIM Community	CPU (Hours)	Connect (Hours)	File Space (**) (Pages)
1) ACT Project "Acquisition of Cognitive Procedures" John R. Anderson, Ph.D. Carnegie-Mellon Univ. NSF IST-80-15357 2/81-2/84 $186,000	0.37	33.88	2866
2) CADUCEUS "Clinical Decision Systems Research Resource" Jack D. Myers, M.D. Harry E. Pople, Jr., Ph.D. University of Pittsburgh NIH RR-01101-07 7/80-6/85 $1,607,717 7/83-6/84 $369,484 NLM LM03710-04 7/80-6/85 $817,884 7/83-6/84 $196,710 NLM New Invest LM03889-02 Gordon E. Banks, M.D. 4/82-3/85 $107,675 4/83-3/84 $35,975 4/84-3/85 $35,975	58.15	895.52	6852
3) CLIPR Project "Hierarchical Models of Human Cognition" Walter Kintsch, Ph.D. Peter G. Polson, Ph.D. University of Colorado NIMH MH-15872-14-16 (Kintsch) 7/81-6/84 $281,085 7/83-6/84 $69,878 NSF (Kintsch) 8/83-7/86 $200,000 IBM (Polson) David Kieras University of Arizona 1/82-12/84 $364,000 1/84-12/84 $145,000	1.38	209.34	750

4) PUFF-VM Project "Biomedical Knowledge Engineering in Clinical Medicine" John J. Osborn, M.D. Med. Research Inst., San Francisco Edward H. Shortliffe, M.D.,Ph.D. Stanford University Johnson & Johnson 1 year $50,000 (*)	0.65	61.20	303
5) SECS Project "Simulation & Evaluation of Chemical Synthesis" W. Todd Wipke, Ph.D. U. California, Santa Cruz NIHEHS ES02845-02 4/82-3/85 $257,801 4/84-3/85 $89,140 Evans & Sutherland Corp. Equipment gift Value $95,000 Stauffer Chemical Co. $6,000	264.61	9877.34	10500
6) SOLVER Project "Problem Solving Expertise" Paul E. Johnson, Ph.D. William B. Thompson, Ph.D. Control Data Corp. (Johnson) 1983-85 $90,000 Microelect. and Info. Ctr. Univ. of MN (Plus Two Colleagues) 1984-1987 $800,000	5.76	356.23	492

7) *** [Rutgers-AIM] ***			
Rutgers Research Resource "Computers in Biomedicine" Saul Amarel, D.Sc. Casimir Kulikowski, Ph.D. Sholom Weiss, Ph.D Rutgers U., New Brunswick NIH RR-00643-12 (Amarel, Kulikowski) 12/82-11/83 $405,304 NIH RR-02230-01 (Kulikowski, Weiss) 12/83-11/87 $3,198,075 12/83-11/84 $989,276	0.52	38.59	1117
8) AIM Pilot Projects	65.85	2227.48	2461
9) AIM Administration	.93	118.75	686
10) AIM Users	57.36	3836.19	9649
	---	---	---
Community Totals	455.56	17654.52	35676

Stanford Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) AGE Project (Core) "Attempt to Generalize" Edward A. Feigenbaum, Ph.D. Dept. Computer Science ARPA MDA903-80-C-0107 (***) (partial support)	11.80	845.30	4076
2) AI Handbook Project (Core) Edward A. Feigenbaum, Ph.D. Dept. Computer Science ARPA MDA903-80-C-0107 (**) (partial support)	11.03	980.94	4425
3) DENDRAL Project "Resource Related Research: Computers in Chemistry" Carl Djerassi, Ph.D. Dennis H. Smith, Ph.D. Dept. Chemistry NIH RR-00612-13 5/82-4/83 $170,710	3.72	183.81	2980
4) EXPEX Project "Expert Explanation" Edward H. Shortliffe, M.D.,Ph.D. Dept. Medicine ONR NR 049-479 1/81-12/83 $456,622 ONR NR049-479 Michael Genesereth 1/84-12/86 $312,070 NSF IST83-12148 Bruce G. Buchanan 3/84-2/87 $330,000 (*) 3/84-2/85 $99,410 (*)	53.75	2391.40	4920
5) GUIDON-NEOMYCIN Project "Exploration of Tutoring & Problem-solving Strategies" Bruce G. Buchanan, Ph.D. William J. Clancey, Ph.D. Dept. Computer Science ONR/ARI N00014-79-C-0302 3/79-3/85 $683,892	45.44	4418.68	5967

6) MOLGEN Project "Applications of Artificial Intelligence to Molecular Biology" Edward A. Feigenbaum, Ph.D. Peter Friedland, Ph.D. Charles Yanofsky, Ph.D. Depts. Computer Science/ Biology NSF MCS-8310236 (Feigenbaum, Yanofsky) 11/83-10/84 $139,215 (*)	106.92	7734.34	10448
7) ONCOCIN Project "Knowledge Engineering for Med. Consultation" Edward H. Shortliffe, M.D.,Ph.D. Dept. Medicine NLM LM-03395 (Shortliffe/ONCOCIN) Edward A. Feigenbaum, Ph.D. 7/79-6/84 $497,420 7/83-6/84 $95,424 NLM LM-00048 7/79-6/84 $196,425 7/83-6/84 $39,502 ONR NR 049-479 1/81-12/83 $456,622 (*) NIH RR-01613 7/83-6/86 $624,455 7/83-6/84 $220,371 NLM LM-04136 8/83-7/86 $211,851 8/83-7/84 $60,517 H.J. Kaiser Family Fdn. 7/83-6/86 $150,000 7/83-6/84 $50,000 ONR N00014-81-K-0004 Michael R. Genesereth (Shortliffe) 1/84-12/86 $512,070 (*) NSF IST83-12148 Bruce G. Buchanan (Shortliffe) 3/84-2/87 $330,000 (*) 3/84-2/85 $99,410 (*)	239.97	14404.62	14389
8) PROTEIN Project "Heuristic Comp. Applied to Prot. Crystallog." Edward A. Feigenbaum, Ph.D. Dept. Computer Science NSF MCS-81-17330 1/82-1/83 $28,976	4.79	635.43	1296

9) RADIX Project "Deriving Medical Knowledge from Time-Oriented Clinical Databases" Robert L. Blum, M.D. Gio C.M. Wiederhold, Ph.D. Depts. Computer Science/ Electrical Engrg. NSF IST-8317858 (Blum) 3/84-3/86 $89597 (*) NLM (Wiederhold) 5/84-11/86 $291,192	79.44	3140.27	8777
10) Stanford Pilot Projects	61.55	4115.02	6097
11) HPP Core AI Research	383.07	29073.96	42202
12) HPP Associates	57.37	1600.31	2997
13) Stanford Associates	27.01	1016.59	1681
14) Medical Information Sciences	5.62	1315.64	587
Community Totals	1091.46	71856.29	110842

SUMEX Staff	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Staff	288.21	17591.82	23292
2) System Associates	16.65	1983.43	7847
Community Totals	304.87	19575.25	31139

System Operations	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Operations	530.54	67375.43	167863
Resource Totals	2382.43	176461.50	345520

(*) Award includes indirect costs.

(**) Supported by a larger ARPA contract MDA-903-80-C-0107 awarded to the Stanford Computer Science Department:

System Reliability

System reliability for the DECsystem 2060 has been much better than with our previous KI10 system. We have had very few periods of particular hardware or software problems. The data below covers the entire period in which the SUMEX-AIM community has used the 2060. The actual downtime was rounded to the nearest hour.

7	18	1
Feb	Mar	Apr

Table 1 : System Downtime Hours per Month - February 83 through Apr 83

11	11	1	2	6	0	11	15	26	13	16	28
May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr

Table 2 : System Downtime Hours per Month - May 83 through Apr 84

Reporting period	:	462 days, 23 hours, 41 minutes, and 42 seconds
Total Up Time	:	454 days, 5 hours, 16 minutes, and 57 seconds
PM Downtime	:	1 days, 14 hours, 2 minutes, and 55 seconds
Actual Downtime	:	7 days, 4 hours, 21 minutes, and 50 seconds
Total Downtime	:	8 days, 18 hours, 24 minutes, and 45 seconds
Mtbf	:	2 days, 16 hours, 30 minutes, and 16 seconds
Uptime Percentage	:	98.45

Network Usage Statistics

The plots in Figure 13 and Figure 14 show the monthly network terminal connect time for the TYMNET and the INTERNET usage. The INTERNET is a broader term for what was previously referred to as Arpanet usage. Since many vendors now support the INTERNET protocols (IP/TCP) in addition to the Arpanet, which converted to IP/TCP in January of 1983, it is no longer possible to distinguish between Arpanet usage and Internet usage on our 2060 system.

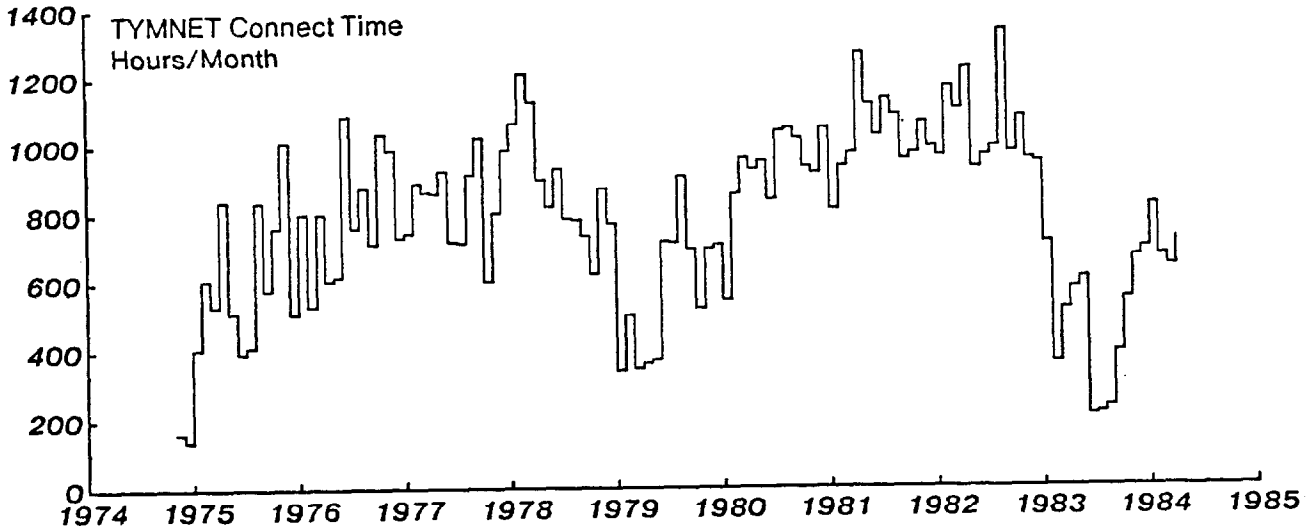

Figure 13: TYMNET Terminal Connect Time

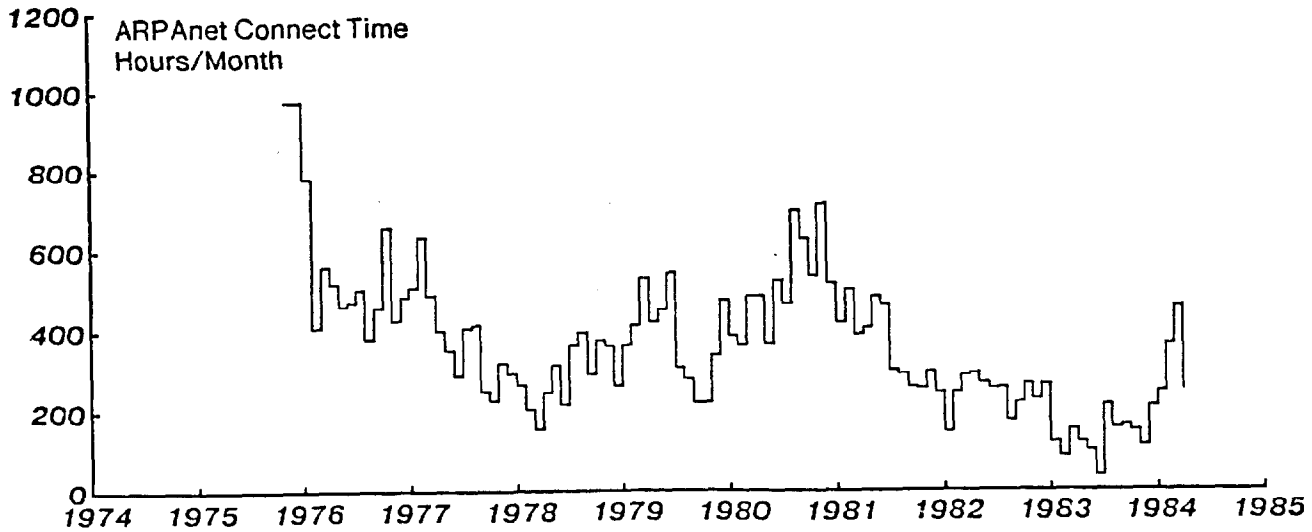

Figure 14: ARPANET Terminal Connect Time

I.A.2.8. SUMEX Staff Publications

The following are publications for the SUMEX staff and include papers describing the SUMEX-AIM resource and on-going research as well as documentation of system and program developments. Many of the publications documenting SUMEX-AIM community research are from the individual collaborative projects and are detailed in their respective reports (see Section II on page 69). Publications for the AGE and AI Handbook core research projects are given there.

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.* IN P. Lykos (Ed.), COMPUTER NETWORKING AND CHEMISTRY, ACS Symposium Series, No. 19, 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.* Proc. DECUS Symposium 2(4), Spring, 1976.

4. Wilcox, C.R.: *The MAINSAIL Project: Developing Tools for Software Portability.* Proc. SCAMC, October, 1977, pp. 76-83.

5. Lederberg, J.L.: *Digital Communications and the Conduct of Science: The New Literacy.* Proc. IEEE 66(11), November, 1978.

6. Wilcox, C.R., Jirak, G.A., and Dageforde, M.L.: *MAINSAIL - Language Manual.* Stanford University Computer Science Report STAN-CS-80-791, 1980.

7. Wilcox, C.R., Jirak, G.A., and Dageforde, M.L.: *MAINSAIL - Implementation Overview.* Stanford University Computer Science Report STAN-CS-80-792, 1980.

In addition, a substantial continuing effort has gone into developing, upgrading, and extending documentation about the SUMEX-AIM resource. These efforts include user guides, help files, and introductory notes, an ARPANET Resource Handbook entry, and policy guidelines.

I.A.2.9. Future Plans

Our plans for the next grant year are based on the Council-approved plans for our 5-year renewal that began in August, 1980. In addition to the specific plans for the next grant year, we include a summary of the overall objectives for this 5-year period to serve as a background. Near- and long-term objectives and plans for individual collaborative projects are discussed in Section II beginning on page 69.

Overall Goals

The goals of the SUMEX-AIM resource are long-term in supporting basic research in artificial intelligence, applying these techniques to a broad range of biomedical problems, experimenting with communication technologies to promote scientific interchange, and developing better tools and facilities to carry on this research. Just as the tone of our renewal proposal derives from the continuing long-term research objectives of the SUMEX-AIM community, our approach derives from the methods and philosophy already established for the resource. We will continue to develop useful knowledge-based software tools for biomedical research based on innovative, yet accessible computing technologies.

For us it is important to make systems that work and are exportable. Hence, our approach is to integrate available state-of-the-art hardware technology as a basis for the underlying software research and development necessary to support the AI work. SUMEX-AIM will retain its broad community orientation in choosing and implementing its resources. We will draw upon the expertise of on-going research efforts where possible and build on these where extensions or innovations are necessary. This orientation has proved to be an effective way to build the current facility and community.

We have built ties to a broad computer science community; have brought the results of their work to the AIM users; and have exported results of our own work. This broader community is particularly active in developing technological tools in the form of new machine architectures, language support, and interactive modalities.

Toward a More Distributed Resource

The initial model for SUMEX as a centralized resource was based on the high cost of powerful computing facilities, which were not readily duplicated. This role is evolving, though, with the introduction of more compact and inexpensive computing technology. Our future goals are guided by community needs for more computing capacity and improved tools to build more effective expert systems, and to test operational versions of AI programs in real-world settings. In order to meet these needs, we must take advantage of a range of newly-developing machine architectures and systems. As a result, SUMEX-AIM will become a more distributed community resource with heterogeneous computing facilities tethered to each other through communications media. Many of these machines will be located physically near the projects or biomedical scientists using them.

The Continuing Role of SUMEX-Central

Even with more distributed computing resources, the central resource will continue to play an important role as a communications crossroad, as a research group devoted to integrating the new software and hardware technologies to meet the needs of medical AI applications, as a spawning ground for new application projects, and as a base for local AI projects. A key challenge will be to maintain the scientific community ties that grew naturally out of the previous colocation within a central facility.

Summary of Five-year Objectives

The long-term objectives of the SUMEX-AIM resource nucleus during the follow-on 5 year period (of which we are in the third year) are summarized below. These are broken into three categories: ***resource operations***, ***training and education***, and ***core research***.

Resource Operations

1. Maintain the vitality of the AIM community -- We will continue to encourage and explore new applications of AI to biomedical research and improve mechanisms for inter- and intra-group collaborations and communications. While AI is our defining theme, we may entertain exceptional applications justified by some other unique feature of SUMEX-AIM essential for important biomedical research. To minimize administrative barriers to the community-oriented goals of SUMEX-AIM and to direct our resources toward purely scientific goals, we plan to retain the current user funding arrangements for projects working on SUMEX facilities. User projects will fund their own manpower and local needs; will actively contribute their special expertise to the SUMEX-AIM community; and will receive an allocation of computing resources under the control of the AIM management committees. There will be no "fee for service" charges for community members. We also will continue to exploit community expertise and sharing in software development, and to facilitate more effective information-sharing among projects.

2. Provide effective computational support for AIM community goals -- We will continue to expand support for artificial intelligence research and new applications work, to develop new computational tools to support more mature projects, and to facilitate testing and research dissemination of nearly operational programs. We will continue to operate and develop the existing central facility as the nucleus of the resource. We will acquire additional equipment to meet developing community needs for more capacity, larger program address spaces, and improved interactive facilities. New computing hardware technologies becoming available now and in the next few years will play a key role in these developments, and we expect to take the lead in this community for adapting these new tools to biomedical AI needs.

3. Provide effective and geographically accessible communication facilities to the SUMEX-AIM community for effective remote collaborations, communications among distributed computing nodes, and experimental testing of AI programs -- We will retain the current ARPANET and TYMNET connections for at least the near-term and will actively explore other advantageous connections to new communications networks and to dedicated links.

Training and Education

1. Assist new and established projects in the effective use of the SUMEX-AIM resource -- Collaborative projects continue to be responsible for the development and dissemination of their own AI programs, but the resource staff will provide general support and will work to make resource goals and AI systems known and available to appropriate biomedical scientists. We will continue to exploit particular areas of expertise within the community for developing pilot efforts in new application areas.

2. Continue to allocate "collaborative linkage" funds to qualifying new and pilot projects to provide for communications and terminal support pending formal approval and funding of their projects -- These funds are allocated in cooperation with the AIM Executive Committee reviews of prospective user projects.

3. Continue to support workshop activities including collaboration with the Rutgers Computers in Biomedicine resource on the AIM Community Workshop and with individual projects for more specialized workshops covering specific application areas or program dissemination

Core Research

1. Continue to explore basic Artificial Intelligence research issues for knowledge acquisition, representation, and utilization; reasoning in the presence of uncertainty; strategy planning; and explanations of reasoning pathways with particular emphasis on biomedical applications -- SUMEX core research funding is complementary to similar funding from other agencies and contributes to the long-standing interdisciplinary effort at Stanford in basic AI research and expert system design. We expect this work to provide the foundation for increasingly effective consultative programs in medicine and for more practical adaptations of this work within emerging microelectronic technologies.

2. Support community efforts to organize and generalize AI tools that have been developed in the context of individual application projects -- This will include work to organize the present state-of-the-art in AI techniques through the development of practical software packages for the acquisition, representation, and utilization of knowledge in AI programs. The objective is to evolve a body of software tools that can be used to more easily build future knowledge-based systems and explore other bionmedical AI applications.

Specific Plans for Year 12

Specific plans for the next grant year (12) are summarized in the paragraphs below. The directions and background for much of this work were given in earlier progress report sections and are not repeated in detail here.

Professional Workstations

We see our major development efforts in year 12 to be in the area of professional workstations, and specifically, to fine tune the integration of these workstations into our networking environment. This involves software integration, support of network protocols, general access to network printing facilities, telnet access to Lisp machines, and overall workstation maintenance and support.

We will also continue to explore the use of low cost workstations within our environment, both as distributed processors for text editing and electronic mail, and as powerful graphic terminals for use with sophisticated programs running on our mainframes. We also see the use of virtual graphics interfaces running on remote workstations to be of continued importance to our progress in the future.

Continued Operation of Existing Hardware

The current SUMEX-AIM facilities represent a large existing investment. We plan to continue development of our main timesharing machine, the DEC2060/TOPS-20 system, and the SUMEX-AIM file server (SAFE), and make changes as necessary to improve the performance of these machines. We do not propose any substantial changes to the other hardware systems (2020, shared VAX, and Lisp Machines). We expect them to continue to provide effective community support and serve as a nucleus for our distributed resource.

Communication Networks

Networks have been centrally important to the research goals of SUMEX-AIM and will become more so in the context of increasingly distributed computing. Communication will be crucial to maintain community scientific contacts, to facilitate shared system and software maintenance based on regional expertise, to allow necessary information flow and access at all levels, and to meet the technical requirements of shared equipment.

We have had reasonable success at meeting the geographical needs of the community during the early phases of SUMEX-AIM through our ARPANET and TYMNET connections. These have allowed users from many locations within the United States and abroad to gain terminal access to the AIM resources and through ARPANET links to communicate much more voluminous file information. Since many of our users do not have ARPANET access privileges for technical or administrative reasons, a key problem impeding remote use has been the limited communications facilities (speed, file transfer, and terminal handling) offered currently by commercial networks. Commercial improvements are slow in coming but may be expected to solve the file transfer problem in the next few years. A number of vendors (AT&T, IBM, XEROX, etc.) have yet to announce commercially-available facilities, but TELENET is actively working in this direction. We plan to continue experimenting with improved facilities as offered by commercial or government sources in the next grant term. We have budgeted for continued TYMNET service and an additional amount annually for experimental network connections.

High-speed interactive terminal support will continue to be a problem since one cannot expect to serve 1200 to 9600 baud terminals effectively over shared long-distance trunk lines with gross capacities of only 9600 to 19200 baud. We feel this is a problem

that is best solved by distributed machines able to effectively support terminal interactions locally and coupled to other AIM machines and facilities through network or telephone links. As new machine resources are introduced into the community, we will allocate budgeted funds with Executive Committee advice to assure effective communications links.

Resource Software

We will continue to maintain the existing system, language, and utility support software on our systems at the most current release levels, including up-to-date documentation. We also will be extending the facilities available to users where appropriate, drawing upon other community developments where possible. We rely heavily on the needs of the user community to direct system software development efforts.

Within the AIM community we expect to serve as a center for software-sharing between various distributed computing nodes. This will include contributing locally-developed programs, distributing those derived from elsewhere in the community, maintaining up-to-date information on subsystems available, and assisting in software maintenance.

Community Management

We plan to retain the current management structure that has worked so well in the past. We will continue to work closely 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 expect the Executive and Advisory Committees to play a continuing role in advising on priorities for facility evolution and on-going community development planning in addition to their recruitment efforts. The composition of the Executive Committee will continue to represent major user groups and medical and computer science applications areas. The Advisory Group membership spans both medical and computer science research expertise. We expect to maintain this policy.

We will continue to make information available about the various projects both inside and outside of the community and, thereby, promote the kinds of exchanges exemplified earlier and made possible by network facilities.

The AIM workshops under the Rutgers resource have served a valuable function in bringing community members and prospective users together. We will continue to support this effort. In July 1984, the AIM workshop will be hosted by Ohio State University. We will continue to assist community participation and provide a computing base for workshop demonstrations and communications. We also will assist individual projects in organizing more specialized workshops as we have done for the DENDRAL and AGE projects.

We plan to continue indefinitely our present policy of non-monetary allocation control. We recognize, of course, that this accentuates our responsibility for the careful selection of projects with high scientific and community merit.

Training and Education Plans

We have an on-going commitment, within the constraints of our staff size, to provide effective user assistance, to maintain high-quality documentation of the evolving software support on the SUMEX-AIM system, and to provide software help facilities such as the HELP and Bulletin Board systems. These latter aids are an effective way to assist resource users in keeping informed about system and community developments and solving usage problems. We plan to take an active role in encouraging the development

and dissemination of community knowledge resources such as the AI Handbook, up-to-date bibliographic sources, and developing knowledge bases. Since much of our community is geographically remote from our machine, these on-line aids are indispensable for self-help. We will continue to provide on-line personal assistance to users within the capacity of available staff through the MM and TALK facilities.

We budget funds to continue the "collaborative linkage" support initiated during the first term of the SUMEX-AIM grant. These funds are allocated under Executive Committee authorization for terminal and communications support to help get new users and pilot projects started.

Core Research Plans

Several systems have been developed in recent years to serve as vehicles for knowledge engineering and research on knowledge representation and its use. Knowledge acquisition (including machine learning) and advanced architectures for AI will be the two areas of most new activity in the coming year. Research on these topics obviously must draw on on-going work in representation and control.

In particular, we will focus on

- Inductive learning of MYCIN-like rules from case data in the domain of diagnosing disorders where the chief complaint is jaundice;
- Learning from experience in domains where the means for interpreting new data are largely contained in the emerging (and thus incomplete and not wholly correct) theory;
- Learning by watching a medical expert diagnose cases presented by NEOMYCIN;
- Investigating complex signal understanding systems for ways to exploit and represent concurrency with a view toward hardware and software architectures that may be capable of several orders of magnitude improvement in performance.

I.B. Highlights

During the past year, the central SUMEX machine has continued to demonstrate its important function as a "seed" environment for new investigators who are embarking on the initial stages of AIM research efforts. SUMEX thus serves as a catalyst and proving ground for new ideas. The potential of such innovations typically needs to be demonstrated in order to provide credible proposals for independent research funding. As more mature projects increasingly turn to professional workstations for their implementation and refinement, we see SUMEX's role as a source of "seed" support for new efforts as being a particularly key element in its function.

In this section we describe several of the highlights of the last year's activities. These include some older projects that have passed important milestones, new pilot projects that have showed remarkable progress in their initial stages, and some other special activities that reflect the impact and influence that SUMEX is demonstrating in the scientific and educational communities.

I.B.1. Progress Towards a Distributed SUMEX-AIM

This past year saw several technical developments at SUMEX which further demonstrate our ability and direction towards establishing SUMEX-AIM as a true distributed resource.

The SUMEX technical staff successfully completed the establishment of a remote computing facility for the Heuristic Programming Project. This new facility, located at 701 Welch Road just off of the Stanford Campus, is connected to SUMEX-AIM via a special 'twisted-pair' ethernet, designed by Nick Veizades, our Senior Electronics Engineer. This new facility also incorporates both 3 and 10 megabit/sec ethernets. The support of these two networks, along with the special ethernet link, necessitated a great deal of work in network software to accommodate this configuration. The resulting technology provides AIM researchers on Welch Road with high speed access to the SUMEX-AIM computer resource despite their remote location. This capability will be of heightened importance when the SUMEX and ONCOCIN groups join the HPP on Welch Road in new quarters sometime during the next year.

One of the most exciting computing prospects for the coming decade is the development of professional workstations. As we have discussed in prior reports, these machines may have a profound impact on biomedicine by serving as the vehicle for the practical export of expert advice systems into the hands of physicians, chemists, biologists, engineers, or other users. SUMEX has continued its investment and research into the use of workstations for biomedical AI research, and the integration of these workstations into a reliable and robust networking environment. In addition to high speed Lisp-based scientific workstations, we believe the use of low cost workstations, which offer suitable local processing power, high resolution screens with easy to use user interfaces, and networking and communications abilities, are vital to the future of our resource.

I.B.2. New Molgen Directions

For several years, the MOLGEN project has focused on research into the applications of symbolic computation and inference to the field of molecular biology. This has taken the specific form of systems which provide assistance to the experimental scientist in various tasks, the most important of which have been the design of complex experiment plans and the analysis of nucleic acid sequences. MOLGEN is now moving into a new phase of research which explores the methodologies scientists use to modify, extend, and test theories of genetic regulation, and then to emulate that process within a computational system.

The first goal of the new work in scientific theory discovery was to study extensively an existing example of the process. Professor Charles Yanofsky's work in elucidating the structure and function of regulation in the ***trp*** operon of E. coli provided an excellent subject that spanned twelve years of research, dozens of collaborators, and almost one hundred research papers.

Extensive interviews have been conducted with Professor Yanofsky and many of his former students and collaborators, and there has been a thorough examination of most of the relevant research papers. This has provided the MOLGEN team with a good understanding of the three major classes of knowledge that were important in the discovery of the theory of regulation in the ***trp*** operon: knowledge about the relevant biological objects, knowledge about the techniques used to elicit new information, and discovery heuristics used to build new models. The major stages in the discovery process have been mapped out, and work has begun on constructing a knowledge base that will represent the state of the world at the beginning of the ***trp*** operon research.

I.B.3. ONCOCIN - An Oncology Chemotherapy Advisor

The ONCOCIN Project, now in its fifth year, and is one of many Stanford research programs devoted to the development of knowledge-based expert systems for application to medicine and the allied sciences. The program is designed to give advice regarding the management of patients receiving cancer chemotherapy. The central issue in this work has been to develop a program that can provide advice similar in quality to that given by human experts, and to insure that the system is easy to use and acceptable to physicians. The work seeks to improve the interactive process, both for the developer of a knowledge-based system, and for the intended end user. In addition, the ONCOCIN group has emphasized clinical implementation of the developing tool so that they can ascertain the effectiveness of the program's interactive capabilities when it is used by physicians who are caring for patients and are uninvolved in the computer-based research activity. ONCOCIN is the first AIM program to have achieved routine (albeit experimental) use by non-collaborating physicians.

ONCOCIN has been used routinely in the Stanford Oncology Clinic for almost three years. Thus, much of the emphasis of this research has been on human engineering so that the physicians will accept the program as a useful adjunct to their patient care activities. The research team has pressed their effort to adapt ONCOCIN to run on professional workstations (specifically the Xerox 1108 "Dandelion") which can eventually be dedicated to full time clinic use. In keeping with other SUMEX experiments in the use of professional workstations as vehicles for implementing medical advice systems, the ONCOCIN team envisions such machines as the model for eventual non-Stanford dissemination of this kind of technology. They have been granted supplemental funding from DRR for three years to support workstation development (along with knowledge base development). They are planning to add all of the protocols in use at the Stanford oncology clinic to ONCOCIN. Major accomplishments in the past year have included the completion of formal studies to evaluate the system's impact in the oncology clinic, the development of a protocol entry system (OPAL) for use by oncologists entering new chemotherapy information into the program, and the development of an 1108 Dandelion environment that is customized for the specialized development needs of this large multi-person project.

I.B.4. New Pilot Projects

This past year saw the addition of several new SUMEX Pilot Projects. Among them are:

PATHFINDER

THE Pathfinder project is directed by Dr. Bharat Nathwani of the Department of Anatomical Pathology, City of Hope National Medical Center, Duarte, California and Dr. Lawrence M. Fagan, Department of Medicine, Stanford University. This project addresses difficulties in the diagnosis of lymph node pathology. Five studies from cooperative oncology groups have documented that, while experts show good agreement with one another, the diagnosis made by practicing pathologists may have to be changed by expert hematopathologists in as many as 50% of the cases. Precise diagnoses are crucial for the determination of optimal treatment. To make the knowledge and diagnostic reasoning capabilities of experts available to the practicing pathologist, The PATHFINDER team has developed a pilot computer-based diagnostic advice system. The project is a collaborative effort of the City of Hope National Medical Center and the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 45 common benign and malignant diseases of the lymph node based on 77 histologic features. The group's research plan, which led to a research proposal to the NIH that is now under consideration, is to develop a full-scale version of the computer program by substantially increasing the quantity and quality of knowledge. They will also further develop techniques for knowledge representation and manipulation appropriate to this application area. The design of the program has been strongly influenced by the INTERNIST/CADUCEUS program that has also been developed on the SUMEX resource. An eventual goal is to merge the diagnostic capabilities of PATHFINDER with a microscope automation effort that Dr. Nathwani is pursuing in collaboration with experts on image processing at Carnegie Mellon University.

Protean

The PROTEAN project involves Dr. Oleg Jardetzky of Stanford Medical School's Nuclear Magnetic Resonance Lab and Prof. Bruce Buchanan of the Computer Science Department. This project has two goals: (a) to use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallizing proteins), and (b) to use protein structure determination as a test problem for experiments with the AI control structure known as the Blackboard Model.

RXDX

The RXDX project is staffed by Dr. Robert Lindsay, Dr. Michael Feinberg, and Dr. Manfred Kochen from the University of Michigan and Dr. Jon Heiser, of the Metropolitan State Hospital in Norwalk, California. This project is developing a prototype expert system to act as a consultant in the diagnosis and management of depression. Health professionals will interact with the program as they might with a human consultant, describing the patient, receiving advice, and asking the consultant about the rationale for each recommendation. The initial prototype is using a knowledge base constructed by encoding the clinical expertise of a skilled psychiatrist in a set of rules. However, the researchers are identifying issues not well addressed by existing rule-based system-building tools (such as EMYCIN) and are anticipating considerable new research in the development of novel techniques for handling such problems.

MENTOR

The MENTOR project is directed by Dr. Stuart M. Speedie and Dr. Terrence F. Blaschke. Dr. Blaschke is Chief of the Division of Clinical Pharmacology in Stanford's Department of Medicine, and Dr. Speedie is a visiting scientist with the Division.

The goal of the MENTOR (Medical EvaluatioN of Therapeutic ORders) project is to design and develop an expert system for monitoring drug therapy for hospitalized patients that will provide appropriate advice to physicians concerning the existence and management of adverse drug reactions. The computer as a recording-keeping device is becoming increasingly common in hospital-based health care, but much of its potential remains unrealized. Furthermore, this information is provided to the physician in the form of raw data which is often difficult to interpret. The wealth of raw data may effectively hide important information about the patient from the physician. This is particularly true with respect to adverse reactions to drugs which can only be detected by simultaneous examinations of several different types of data including drug data, laboratory tests and clinical signs.

I.B.5. Major Books on Medical Artificial Intelligence

Just as the well known *Handbook of Artificial Intelligence* was developed on SUMEX several years ago, the resource has served as the focus for the development of two new books that are being published in 1984. Each book describes research projects that were largely dependent upon the SUMEX-AIM network for their successful implementation. Bruce Buchanan and Ted Shortliffe have edited a large collection of papers regarding the MYCIN system and its derivatives. They have also written new material and analyzed the results of the decade's experiments. The resulting volume, titled *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, will be published by Addison-Wesley in June.

A second volume, to be published by Addison-Wesley in July, is a collection of papers on AIM research efforts. The book, entitled *Readings in Medical Artificial Intelligence: the First Decade*, was edited by Bill Clancey and Ted Shortliffe. Its 21 chapters summarize much of the research that SUMEX has helped spawn.

I.B.6. Training in Medical Information Science

Stanford's nascent program in Medical Information Sciences, mentioned briefly in last year's annual report, has matured significantly in the past 12 months. There will be 9 trainees in the program in September 1984, 7 working towards PhD degrees and 2 towards the MS degree. Of these trainees, 7 have MD degrees or are concurrently enrolled as medical students. Two of the trainees are playing central roles in the PATHFINDER research mentioned above, and several others are involved in ongoing AIM research using SUMEX facilities. The program has been awarded post-doctoral training support from the National Library of Medicine, received an equipment gift of four 9836 workstations from Hewlett Packard Company, and has received additional industrial and foundation grants for student support. We believe that SUMEX has been an important element in the rich medical computing research environment at Stanford that has in turn led to the successful implementation of this novel training effort. It is our belief that the medical computing and AIM communities, as well as biomedicine in general, will benefit greatly from an increased number of people trained to undertake research at the interface between medicine and computer science.

I.C. Administrative Changes

Carole Miller, who had served as the Administrative Assistant for SUMEX since 1974, accepted a new position as the Administrative Assistant of the Heuristic Programming Project in August of 1983. Carole has since moved on to become the Administrative Services Manager for the Center for Research on International Studies here at Stanford.

Patricia (Patti) M. McCabe has succeeded Carole as the Administrative Assistant for SUMEX-AIM. Patti comes to SUMEX-AIM from the Sponsored Projects Office at Stanford University where she was responsible for contracts and grant management, and was the primary liaison between Stanford University and the National Institutes of Health.

Roy Maffly stepped down as the SUMEX-AIM Liaison to devote more time to his responsibilities within the Stanford Medical Center. Larry Fagan, who returned to Stanford this past year as a Senior Research Associate in the Department of Medicine, has taken over for Roy as the new SUMEX AIM liaison.

I.D. Resource Management and Allocation

The mission of SUMEX-AIM, locally and nationally, entails both the recruitment of appropriate research projects interested in medical AI applications and the catalysts of interactions among these groups and the broader medical community. These user projects are separately-funded and autonomous in their management. They are selected for access to SUMEX on the basis of their computer and biomedical scientific merits, as well as their commitment to the community goals of SUMEX. Currently active projects span a broad range of applications areas such as clinical diagnostic consultation, molecular biochemistry, molecular genetics, medical decision making, and instrument data interpretation (Descriptions of the individual collaborative projects are in Section II beginning on page 69).

I.D.1. Management Committees

Since the SUMEX-AIM project is a multilateral undertaking by its very nature, several management committees have been created 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, Prof. Feigenbaum and the 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 A.

For the national community, two committees serve complementary functions. An *Executive Committee* oversee's the operations of the resource as related to national users and renders final decisions on authorizing admission for new projects and revalidating continued access for existing 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, significant new development projects, etc.). The Executive Committee oversees the planning and implementation of the AIM Workshop series, and assures coordination with other AIM activities as well. The Committee will continue to 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 A.

The Executive Committee met in 1983 during the AIM Workshop and via teleconferencing sessions. Items addressed during the committee meetings were final decisions on admissions of new AIM pilot projects, and the annual re-evaluation of continued access for AIM projects. In the latter area, a decision was reached after long and careful review to phase the SECS project out of SUMEX-AIM. The committee was concerned over the system impact of this project versus the current relevance and innovativeness of its research for AI. The implementation of this decision will be to phaseout SECS use in a fair and orderly manner, allowing for reduced system use until the completion of existing project commitments in March, 1985.

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 A.

These committees have actively functioned in support of the resource. Except for 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. Other solicitations of advice requiring review of sizeable 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 will continue to make information available about the various projects both inside and outside of the community and thereby promote the kinds of exchanges exemplified earlier and made possible by network facilities.

I.D.2. New Project Recruiting

We continue to see a very strong interest in Artificial Intelligence applications to medicine. We receive several inquiries a week, stimulated by information on SUMEX-AIM or the SUMEX-AIM subprojects. We are actively recruiting the best of these inquiries as pilot projects to provide new activities to replace projects that have matured and moved off of the SUMEX-AIM machine. A presentation was made at the American Association of Artificial Intelligence conference in August, 1983 to provide general information about SUMEX-AIM and encourage additional users. Additional information about SUMEX-AIM projects is available through well-attended presentations at national conferences in Artificial Intelligence. In addition, interest in the Artificial Intelligence approach to medical decision making has strongly increased in the national medical computing conferences. SUMEX-AIM related researchers are often the key personnel at these presentations.

During the Fall of 1983, two national and two Stanford-related projects were initiated. Many other interested researchers took advantage of SUMEX's ability to allow experimental access to existing computer programs. In addition, some of the more stable software for developing medical applications is now provided on tape for implementation on host computers outside of the SUMEX-AIM environment.

The criteria for the acceptance of new pilot projects continues to concentrate on the potential for excellence, and the novelty of the proposed concepts. We continue to seek projects that will extend our understanding of basic science issues underlying the application of the artificial intelligence approach to medical decision making. Thus, a project that will break new ground will be preferred to a project that uses existing ideas in a new area of medicine. We also encourage pilot projects to collaborate with of the existing bases of expertise in artificial intelligence techniques. Developing a new pilot project now requires more background and understanding of previous work in AI in medicine. However, the time needed to build a first prototype version may be substantially decreased by the use of packages developed by other SUMEX-AIM projects. SUMEX-AIM provides a unique opportunity for the development of pilot projects. We hope to build the number of pilot projects consistent with SUMEX resources and the availability of worthy project proposals.

I.D.3. Stanford Community Building

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 current and state-of-the-art AI concepts and techniques. This project has had enthusiastic support from the students, and the work has culminated in the publication of a three-volume handbook set named the Handbook of AI, published by William Kaufman Press.

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

I.D.4. Existing Project Reviews

We have conducted a continuing careful review of on-going SUMEX-AIM projects to maintain a high scientific quality and relevance to our medical AI goals and to maximize the resources available for newly-developing applications projects. At meetings of the AIM Advisory Group and Executive Committee this past year, all of the national AIM projects were reviewed. These groups recommended continued access for most formal projects on the system, and the phaseout of the SECS project, details of which are covered on page 62.

I.D.5. Resource Allocation Policies

Policies have been established to control the allocation of critical facility resources (file space and central processor time) on the SUMEX-AIM 2060. File space management begins with an allocation of file storage, defined for each authorized project in consultation with the management committees. This allocation for any given project is redistributed among project members as directed by the individual principal investigators. System enforcement of project allocations is done on a weekly basis. As the weekly file dump is done, if the aggregate space in use by a project exceeds its allocation, files are archived from associated user directories which are over allocation until the project is within its authorized limits.

We are using the TOPS-20 class scheduler to attempt to enforce the 40:40:20 balance in terms of CPU utilization and to avoid system and user inefficiencies under overload conditions. In practice, the 40:40 split between Stanford and non-Stanford projects is fairly well realized (see Figure 10 on page 34 and the tables of recent project usage on page 36).

Our job-scheduling controls bias the allocation of CPU time based on per cent time consumed relative to the time allocated according to the 40:40:20 community split. However, the controls are "soft" in that they do not waste computer cycles if users below their allocated percentages are not on the system to consume those cycles. In the early years, the operating disparity in CPU use reflected 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 overload control system reported previously, the Stanford community experienced mutual contentions and delays while the AIM group had relatively open access to the system.

This disparity in usage has disappeared in recent years with the growth of the national user community, and we enabled overload controls for the national community as well. 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, associates, 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 that experienced by other network sites, perhaps because of the personal attention directed by senior staff to 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.

We also have encouraged mature projects to apply for their own machine resources in order to preserve the SUMEX-AIM resource for research and development efforts and to support projects unable to justify their own machines. The Rutgers resource has its own 2060 machine, part of which is allocated for AIM use, and the CADUCEUS project has installed a VAX 11/780 machine to support its planned development and program testing work. Profs. Lesgold and Greeno's "Simulation of Cognitive Processes" Project has moved entirely to their own local VAX.

I.E. Dissemination Efforts

Throughout its existence, SUMEX-AIM has expended substantial effort toward disseminating information about its activities as a resource and about the work of individual collaborative projects. We continue to make many presentations at professional meetings, to provide services to demonstrate developed AI programs to interested groups and individuals, to welcome visitors, and to work in organizing workshops within the SUMEX-AIM community to introduce our research to collaborating professional communities. We also directed considerable effort in the past toward working with the Research Resources Information Center to produce the "Seeds of Artificial Intelligence" monograph and other publications and press articles to address a broader community of technical and lay people.

Software Distribution

SUMEX continues to support various projects in the distribution of versions of their software to requesting individuals or groups. Following is a summary of software dissemination this past year:

EMYCIN — Both the "executable" and "source" versions of the EMYCIN distribution package were restructured for clarity and ease of installation. Thirty copies of the EMYCIN package have been generated for distribution of which about 6 were sources only. An Interlisp-VAX version of EMYCIN is now available, thanks to Ray Bates of USC-ISI, who did the conversion. This runs under UNIX and VMS.

AGE — Twenty-two copies of the AGE system have been distributed. Nearly half of these have been copies requested in ANSI format indicating they were evidently going to non-Tops20 sites (probably Vaxes). As with the EMYCIN system, Ray Bates at USC-ISI has converted AGE to run under Interlisp-VAX. A version is also available for the Xerox 1108 series Lisp workstations.

GENET — In conjunction with the phaseout of the GENET community on SUMEX, a software package comprised of programs and databases developed by researchers at Stanford and elsewhere was assembled for distribution to interested GENET users. Versions of the software were provided for use on both DEC-10 and DEC-20 systems operating under TOPS-10, TENEX, and TOPS-20. Installation procedures were documented, and a substantial amount of telephone consultation was provided. The package has been well-received and appears to be in active use at many of the 21 academic sites to which it was sent. Only one copy of the complete Genet system was set out in the past year. However, several sets of Genet related data files have been distributed. This includes several copies of the NIH and EMBL Sequence Libraries. A limited amount of operations support has been given to Brutlag's interaction with Sam Karlin of the Math department and a variety of other groups.

MRS — Twenty-two copies of MRS have been distributed through Sumex. Several others have also been distributed directly by the HPP. Most have been sent out to VAX/Unix sites or Symbolics Lisp machine sites.

SACON	Two copies of SACON have been prepared and distributed.
GLISP	Two copies of GLISP were distributed.

I.F. Comments on the Biotechnology Resources Program

Resource Organization

We continue to believe that the Biotechnology Resources Program is one of the most effective vehicles for developing and disseminating technological tools for biomedical research. The goals and methods of the program are well-designed to encourage building of the necessary multi-disciplinary groups and merging of appropriate technological and medical disciplines. In our experience with the SUMEX-AIM resource, several elements of this approach seem to emerge as key to the development and management of an effective resource:

1. Effective Management Framework -- There needs to be an explicit agreement between the BRP and the resource principal investigator which establishes a clear mandate for the resource and its allocation, provides worthwhile incentives for the host institution and investigator to invest the necessary substantial professional career time to develop and manage the resource, and ensures equitable distribution of resource services to its target community.

2. Close Working Relationship with the NIH -- A resource is a major and often long-term investment of money and human energy. A close and mutually-supportive working relationship between resource management, its advisory committees, and the NIH administration is essential to assure healthy development of the resource and its relationship to its user community. We at SUMEX-AIM have benefited immensely from such a relationship with Dr. William R. Baker, Jr., in the evolution of the SUMEX-AIM community. We look forward to a continuing mutually beneficial relationship with Dr. Baker's successor at the NIH.

3. Freedom to Explore Resource Potential -- A resource, by its nature, operates at the "cutting edge" in developing its characteristic technology and learning to effectively disseminate it to the biomedical community at large. The BRP should not impose artificial constraints on the resource for commercializing its efforts (fees for service) or developing its potential (funding duration limits or annual budget ceilings). Such artificial policy impositions can serve to undermine the very goals central to the BRP's reason for existence. Satisfactory policies in this regard have been worked out and should be retained.

Electronic Communications

SUMEX-AIM has pioneered in developing more effective methods for facilitating scientific communication. Whereas face-to-face contacts continue to play a key role, in the longer-term we feel that computer-based communications will become increasingly important to the NIH and the biomedical community. We would like to see the BRP take a more active role in promoting these tools within the NIH and its grantee community.

II. Description of Scientific Subprojects

II.A. Scientific Subprojects

The following subsections report on the AIM community of projects and "pilot" efforts including local and national users of the SUMEX-AIM facility at Stanford. However, those projects admitted to the National AIM community which use the Rutgers-AIM resource as their home base are not explicitly reported here.

In addition to these detailed progress reports, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form. However, we have included here briefer summary abstracts of the fully-authorized projects in Appendix B on page 209.

The collaborative project reports 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. Project rationale
- B. Medical relevance and collaboration
- C. Highlights of research progress
  - --Accomplishments this past year
  - --Research in progress
- D. List of relevant publications
- E. Funding support

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE
- A. Medical collaborations and program dissemination via SUMEX
- B. Sharing and interactions with other SUMEX-AIM projects (via computing facilities, workshops, personal contacts, etc.)
- C. Critique of resource management (community facilitation, computer services, communications services, capacity, etc.)

III. RESEARCH PLANS
- A. Project goals and plans
  - --Near-term
  - --Long-range
- B. Justification and requirements for continued SUMEX use
- C. Needs and plans for other computing 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. The only exceptions are the respective lists of relevant publications which have been uniformly formatted for parallel reporting on the Scientific Subproject Form.

II.A.1. 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 Feigenbaum as Principal Investigator.

In addition to the progress reports presented here, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form.

II.A.1.1. EXPEX - Expert Explanation Project

EXPEX - Expert Explanation Project

Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

EXPEX is not a single project but a combination of efforts that are directed at basic issues in the development of representational schemes to facilitate knowledge acquisition and explanation. The work includes not only the study of fundamental representational formalisms but also the encoding of various types of knowledge, such as causal information and user models. In addition, to complement these research directions, the project has served as the focus for preparing three books on medical computing research.

We believe that the productivity of basic computer science research tends to be heightened by experiments that deal with significant real world problem domains. Challenges drawn from chemistry, medicine, and molecular biology have introduced additional complexity to expert systems work at Stanford, but have simultaneously forced system developers to respond to pragmatic constraints and user demands that have had a significant impact on the basic AI techniques selected or developed. Thus, we believe that creative investigation into symbolic reasoning techniques is facilitated by working in real world settings where the application forces us to avoid oversimplification. Much of our research effort therefore deals with medical domains (viz., endocrinology and renal pathophysiology).

B. Medical Relevance and Collaboration

Our interest in explanation derives from the insights we gained in developing explanatory capabilities for the MYCIN system. In the case of MYCIN and its descendents, we have been able to generate intelligible explanations by taking advantage of its rule-based representation scheme. Rules can be translated into English for display to a user, and their interactions can also be explicitly demonstrated. By adding mechanisms for understanding questions expressed in simple English, we were able to create an interactive system that allowed physicians to convince themselves that they agreed with the basis for the program's recommendations. The limitations of the explanations generated in this way have become increasingly obvious, however, and have led to improved characterization of the kinds of explanation capabilities that must be developed if clinical consultation systems are to be accepted by physicians. The potential use of workstation graphics as a means of avoiding natural language issues in the explanation process is also an area of great promise with which we are currently experimenting.

With these motivations in mind, we are involved in a series of research projects that address medical knowledge representation and explanation. The individual projects include the following:

1. Mr. Greg Cooper's NESTOR program uses a detailed knowledge base regarding pathophysiologic relationships in hypercalcemia. The program is designed to critique a physician's hypothesis regarding a proposed explanation for a set of patient manifestations when an elevated serum calcium has been observed. Of particular interest is the techniques Cooper has developed for using knowledge of causality to avoid the assumption of conditional independence commonly used in Bayesian diagnosis systems.

2. Mr. John Kunz has represented the knowledge of renal pathophysiology, including the ***quantitative*** relationships that characterize the way in which the body manages water and electrolytes, to develop a consultation and analysis system (AI/MM) that melds mathematics and AI techniques.

3. Building on his earlier experience with developing an explanation capability for NEOMYCIN (in collaboration with the GUIDON project members as outlined elsewhere in this report), Dr. Glenn Rennels has begun to work on a new system that uses knowledge of medicine to help formulate and resolve complex decision analyses. Convinced that decision analytic techniques would be better accepted in medicine if the physician were to interact with a knowledge-based interface (rather than with the decision trees themselves), Dr. Rennels has made use of "influence diagrams" as a central method for guiding the interaction. The explanation issues become especially evident when an analysis is complete and his system needs to generate a defense for the recommendation it has made.

4. Mr. Curt Langlotz has continued to work on a hypothesis assessment module for the ONCOCIN system. This program uses a ***critiquing model*** which inherently involves advanced explanation techniques. The work uses the Xerox 1108 professional workstation (Dandelion) and is further described in the ONCOCIN Project portion of this annual report.

5. During 1983, Ms. Shoko Tsuji completed a project using the Xerox workstation to experiment with graphical techniques for examining, manipulating, expanding, and editing a large medical knowledge base. Also working in the context of ONCOCIN, her code was designed for use by knowledge engineers. The work has inspired subsequent work in building an interface for the non-programmer clinician who wishes to write and test new protocols in the ONCOCIN environment. The project is described in greater detail in the ONCOCIN portion of the annual report.

To complement these basic research activities, we have prepared two books on Artificial Intelligence in Medicine and are beginning work on a third (see Section C for details).

C. Highlights of Research Progress

C.1 The NESTOR System

NESTOR is intended to allow a user to input patient data plus a hypothesis, and then have the system critique that hypothesis in light of the data. The system, an evolving thesis project that is largely the work of Mr. Greg Cooper, relies on basic associational information drawn in part from the INTERNIST-I knowledge base but supplemented with causal and temporal associations.

The motivation behind this research is the conviction that physicians want active

control of the diagnostic process and that they also want and need a system that explains, in a user-tailored way, its evaluation of the physician's hypothesis. There may be times when the user wants to give complete control to NESTOR and just be in a mode of answering questions, but we feel that this should be an option and not a requirement. It is observations such as these that have also accounted for the hypothesis assessment work underway in the ONCOCIN research, briefly mentioned above and further described in the section of this report dealing with that project.

The initial NESTOR system is now largely complete and is undergoing evaluation at this time. Of particular interest is the adequacy of the techniques developed for allowing NESTOR to avoid the traditional assumption of conditional independence used in Bayesian systems. Also, because NESTOR's probabilistic model is more formal than the ad hoc scheme used in, say, INTERNIST, the assumptions made by our system are more explicit.

We have also developed search techniques that allow NESTOR to explore efficiently a very large search space in order to find the most probable (multiple disease) hypothesis. This technique is general and can be applied to many nonmedical problems where the goal is to find the most probable hypothesis among many possibilities.

C.2 Integrating Mathematical Models with AI Methods

This research project, known as AI/MM, is the dissertation research of Mr. John Kunz. The system integrates AI and simple mathematics to analyze a physiological model. In a selected medical domain (renal physiology), we have built a computer program based on these techniques. It analyzes physiological behavior, diagnoses abnormality, and explains the rationale for its analyses. The program fits data to the model, identifies whether the data are abnormal, and identifies the possible causes and effects of any abnormalities. The physiological model is based on knowledge about anatomy, the behavior of the physiological system, and the mechanism of action of the system. It's validity has been tested by having it analyze many of the problems discussed in Valtin's text *Renal Function*.

The specific aims of this project have been to:

1. Develop a vocabulary for a physiological model. The vocabulary represents the "basic physiology" of a biological system and appears to be adequate to express the concepts included in an introductory professional-level physiology text.

2. Develop a reasoning system which can solve problems expressed in the vocabulary.

3. Demonstrate the basic necessity, appropriateness and limitations of the vocabulary and reasoning procedure.

C.3 Knowledge-Based Explanations in a Decision Analysis Environment

This new project, thesis research by Dr. Glenn Rennels, is motivated by the observation that AI techniques could greatly facilitate a user's effort to specify the details of a complex clinical decision task and to seek assistance with that task. Although decision theoretic notions have been shown to be relevant to such medical problems, they have largely been unused by clinicians, even when computer-based solutions have been offered. We believe that an intelligent system should be able to ***guide*** the definition of the decision task and ***explain*** the results of the analysis without requiring that a user be familiar with the underlying decision analytic techniques being used to solve the problem.

The basic notion is to use directed graphs, termed "influence diagrams" as a language for communication with a physician at a graphical display terminal. Nodes in these graphs are defined by the user who is seeking advice, and their structure and meaning is largely intuitive. The task of converting influence diagrams to decision trees is a knowledge-based problem that is potentially well-suited for a solution that uses AI methods. Similarly, the results of a decision analysis, including the sensitivity analysis, will need to be explained to the physician user in terms of influence diagrams and knowledge of the domain. The necessary knowledge structures are currently being designed, and an early prototype system is operational. The research uses a 9836 workstation donated to the Medical Information Sciences Training Program by Hewlett-Packard Company and soon to be networked to the SUMEX 2060.

C.4 Books on Medical Artificial Intelligence and Medical Computing

We have completed two books, both of which are in press and due to be published in mid-1984:

- Clancey, W.J. and Shortliffe, E.H. ***Readings in Medical Artificial Intelligence: The First Decade.*** Reading, MA: Addison-Wesley, 1984.
- Buchanan, W.J. and Shortliffe, E.H. ***Rule-Based Expert Systems: the MYCIN Experiments of the Stanford Heuristic Programming Project.*** Reading, MA: Addison-Wesley, 1984.

In addition, we have just begun work on a textbook for students beginning to study medical computing and artificial intelligence. This multi-authored volume should be completed in draft form by the end of 1984. A 1985 publication date is contemplated.

- Shortliffe, E.H., Wiederhold, G.C.M., and Fagan, L.M. ***An Introduction to Medical Computer Science.*** Reading, MA: Addison-Wesley (in preparation).

D. Publications Since January 1983

1. Shortliffe, E.H. Medical consultation systems: designing for doctors. In *Designing for Human-Computer Communication* (M.S. Sime and M.J. Coombs, eds.), Chapter 8, pp. 209-238, London: Academic Press, 1983.

2. Shortliffe, E. H. Medical Cybernetics: The Challenges of Clinical Computing. In *Technology International Stability, and Growth*, S. Basheer Ahmed and Alice P. Ahmed, editors; Chapter 12, pp. 148-165; Associated Faculty Press, Inc., Port Washington, New York, 1984.

3. (*) Shortliffe, E.H. and Fagan, L.M. Expert systems research: modeling the medical decision making process. In ***An Integrated Approach to Monitoring*** (J.S. Gravenstein, R.S. Newbower, A.K. Ream, and N.T. Smith, eds.), pp. 183-200, Woburn, MA: Butterworth's, 1983.

4. Duda, R.O. and Shortliffe, E.H. Expert systems research. ***Science***, 220:261-268 (1983).

5. (*) Langlotz, C.P. and Shortliffe, E.H. Adapting a consultation system to critique user plans. ***International Journal of Man-Machine Studies***, 19:479-496 (1983)

6. Shortliffe, E.H. Hypothesis generation in medical consultation systems: artificial intelligence approaches. In *MEDINFO 83* (J.H. van Bemmel, M. Ball, and O. Wigertz, eds.), pp. 480-483, North Holland, Amsterdam, 1983.

7. (*) Tsuji, S. and Shortliffe, E.H. Graphical access to the knowledge base of a medical consultation system. *Proceedings of AAMSI Congress 83*, pp. 551-555, San Francisco, Ca., May 1983.

8. Shortliffe, E.H. The science of biomedical computing. In *Meeting the Challenge: Informatics and Medical Education* (J.C. Pages, A.H. Levy, F. Gremy, and J. Anderson, eds.), pp. 1-10, Amsterdam: North-Holland, 1983. To be reprinted in Medical Informatics, 1984.

9. (*) Kunz, J.C., Shortliffe, E.H., Buchanan, B.G., and Feigenbaum, E.A. Comparison of techniques of computer-assisted decision making in medicine. In *Pure and Applied Biostructure* (Claudio Niccolini, Ed.), Singapore: World Press, 1983.

10. (*) Kunz, J.C., Shortliffe, E.H., Buchanan, B.G., Feigenbaum, E.A. Computer-assisted decision making in medicine. *Journal of Philosophy and Medicine*, Summer 1984 (in press).

11. (*) Hasling, D. W., Clancey, W. J., and Rennels, G. Strategic explanations for a diagnostic consultation system. *International Journal of Man-Machine Studies*, Spring 1984 (in press).

12. Shortliffe, E.H. Reasoning methods in medical consultation systems: artificial intelligence approaches (tutorial). *Computer Programs in Biomedicine*, January 1984 (in press).

E. Funding Support

Grant Title: "The Development of Representation Methods to Facilitate Knowledge Acquisition and Exposition in Expert Systems"
Principal Investigator: Edward H. Shortliffe
Agency: Office of Naval Research; ID Number: NR 049-479
Term: January 1981 to December 1983
Total award: $456,622

Grant Title: "Research on Introspective Systems"
Principal Investigator: Michael R. Genesereth
Agency: Office of Naval Research; ID Number: NR 049-479
Term: January 1984 to December 1986
Total award: $312,070

Grant Title: "Information Structure and Use in Knowledge-Based Expert Systems"
Principal Investigator: Bruce G. Buchanan
Agency: National Science Foundation; ID Number: 83-12148
Term: March 1984 to February 1987
Total award: $300,000 (includes indirect costs)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

None of these new programs is yet ready for dissemination. They are mostly

fundamental research experiments with limited clinical usefulness other than as demonstration projects. Our past experience has shown, however, that SUMEX provides a superb vehicle for demonstrating systems, even at a distance.

The new book writing effort will in particular be facilitated by SUMEX, much as the *AI Handbook* was in the past. A multi-authored text of this type, particularly one for which the authors are spread at numerous different universities around the country, would be a nightmare to compile if it were not for the SUMEX resource. Many of the contributors to the book have been assigned SUMEX accounts for purposes of manuscript preparation. Online manuscript work through the shared facility, coupled with messaging capabilities, will greatly enhance the efficiency and accuracy of the developing chapters and the editing process.

B. Sharing and Interaction with Other SUMEX-AIM Projects

Although our EXPEX work is young, we are already benefiting from interactions with other researchers who use the SUMEX-AIM resource. The NESTOR work in particular has depended on access to the INTERNIST-1 knowledge base and on frequent exchange of messages with the researchers at the University of Pittsburgh. Similarly, our collaboration with the GUIDON research team for the implementation of an explanation capability would not have been possible without the facilitated communication and shared file access available via SUMEX.

C. Critique of Resource Management

SUMEX continues to provide a superb environment for research of this kind. Not only is the 2060 a well managed resource under Ed Pattermann's leadership, but the hypothesis assessment and graphical query systems are dependent upon access to high performance professional workstations, and we are delighted with the resources that SUMEX has provided us in this regard.

III. RESEARCH PLANS

A. Project Goals and Plans

We anticipate completion of many of these basic research efforts during the coming year. Cooper's NESTOR work is largely complete, and a thesis document is anticipated in June 1984. Similarly, Kunz has completed his work on AI/MM, and his dissertation is approaching completion. Both Cooper and Kunz have completed their oral examinations on this work.

The project of Tsuji is complete and she has now left Stanford. However, the code she developed is being modified for ongoing use in the ONCOCIN environment.

The project of Langlotz continues to be an active research effort within the ONCOCIN project. His plan for the coming year is briefly outlined in the ONCOCIN portion of this annual report.

The work of Rennels, which is just getting underway, will be better formulated by next year at this time. We expect the project to last at least two more years, however.

The textbook preparation is scheduled for completion in approximately one year, with publication anticipated during 1985.

B. Requirements for Continued SUMEX Use

All the work we are doing is largely dependent on the SUMEX resource. The new work of Rennels is using Hewlett-Packard 9836 workstations owned by the Medical Information Sciences training program, but Dr. Rennels continues to be dependent upon SUMEX for communication and collaboration. Of the other projects, only the hypothesis assessment and graphical query projects are sufficiently mature to justify their transfer to one of the SUMEX personal workstations, so the new 2060 continues to be a key element in our research plan.

In addition, we have long appreciated the benefits of GUEST and network access to the programs we are developing. SUMEX greatly enhances our ability to obtain feedback from interested physicians and computer scientists around the country. As our programs continue to mature, it will become increasingly important that we be able to make them available for demonstration and for access by distant collaborators via the SUMEX network.

C. Requirements for Additional Computing Resources

The mainframe machine should continue to provide a suitable environment for most of our work in the months ahead. We have no plans to transfer NESTOR, or AI/MM to other hardware soon.

D. Recommendations for Future Community and Resource Development

We are very satisfied with the facilities SUMEX has provided since the upgrade to the DEC 2060. Other than continued acquisition of professional workstations that can be shared by some of the more mature programs in this set of projects, we have no requests for additional acquisitions or resource development at this time.

II.A.1.2. GUIDON/NEOMYCIN Project

GUIDON/NEOMYCIN Project

William J. Clancey, Ph.D.
Department Computer Science
Stanford University

Bruce G. Buchanan, Ph.D.
Computer Science Department
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The GUIDON/NEOMYCIN Project is a research program devoted to the development of a knowledge-based tutoring system for application to medicine. This work derived from our first system, the MYCIN program. That research gave way to three sub-projects (EMYCIN, GUIDON, and ONCOCIN) described in previous annual reports. EMYCIN has been completed and its resources reallocated to other projects. GUIDON and ONCOCIN have become projects in their own right.

The key issue for the GUIDON/NEOMYCIN project is to develop a program that can provide advice similar in quality to that given by human experts, modeling how they structure their knowledge as well as their problem solving procedures. The consultation program using this knowledge is called NEOMYCIN. NEOMYCIN's knowledge base, designed for use in a teaching application, will become the subject material used by a family of instructional programs referred to collectively as GUIDON2. The problem-solving procedures are developed by running test cases through NEOMYCIN and comparing them to expert behavior. Also, we are using NEOMYCIN as a test bed for the explanation capabilities that will eventually be part of our instructional programs.

The purpose of the current contract, now in its sixth of six years, is to construct an intelligent tutoring system that teaches diagnostic strategies explicitly. By strategy, we mean plans for establishing a set of possible diagnoses, focusing on and confirming individual diagnoses, gathering data, and processing new data. The tutorial program will have capabilities to recognize these plans, as well as to articulate strategies in explanations about how to do diagnosis. The strategies represented in the program, modeling techniques, and explanation techniques are wholly separate from the knowledge base, so can be used with many medical (and non-medical) domains. That is, the target program will be able to be tested with other knowledge bases, using system-building tools that we provide.

B. Medical Relevance and Collaboration

There is a growing realization that medical knowledge, originally codified for the purpose of computer-based consultations, may be utilized in additional ways that are medically relevant. Using the knowledge to teach medical students is perhaps foremost among these, and NEOMYCIN continues to focus on methods for augmenting clinical knowledge in order to facilitate its use in a tutorial setting. A particularly important

aspect of this work is the insight that has been gained regarding the need to structure knowledge differently, and in more detail, when it is being used for different purposes (e.g., teaching as opposed to clinical decision making). It was this aspect of the GUIDON research that led to the development of NEOMYCIN, which is an evolving computational model of medical diagnostic reasoning that we hope will enable us to better understand and teach diagnosis to students. An important additional realization is that these structuring methods are beneficial for improving the problem-solving performance of consultation programs, providing more detailed and abstract explanations to consultation users, and making knowledge bases easier to maintain.

As we move from technological development of explanation and student modeling capabilities, we will in the next year begin to collaborate more closely with the medical community to design an effective, useful tutoring program. Stanford Medical School faculty, such as Dr. Maffly, have shown considerable interest in this project. A research fellow associated with Maffly, Curt Kapsner, MD, joined the project last year to serve as medical expert and liaison with medical students at Stanford.

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

C.1.1 The NEOMYCIN Consultation Program

NEOMYCIN is distinguished from other AI consultation programs by its uses of an explicit set of domain-independent meta-rules for controlling all reasoning. These rules constitute the diagnostic procedure that we want to teach to students: the stages of diagnosis, how to focus on new hypotheses, and how to evaluate hypotheses. It has been a major undertaking, separate from the problem of representing disease knowledge, to design and test this diagnostic procedure. Such modifications require changing our conception of how disease knowledge is organized. For example, this year we partitioned disease findings into "non-specific" and "red flag" (those requiring explanation), augmenting the diagnostic procedure to use this information for focusing on hypotheses. A second change is to have the program reason about the disease process more generally. By associating symptoms by organ system, NEOMYCIN now has primitive means to infer when a disease process began. It also makes more complete use of severity, location, and progression information to discriminate among hypotheses.

During this past year, we completely reworked the program's knowledge of non-meningitis cases. This is important if we wish to teach students to consider the competitors of meningitis and how to discriminate among them. The goal is to prepare the program for presenting these (or similar) cases to students. In order to test the modeling component, it is necessary to ensure that the program has sufficient expertise to recognize good student behavior. All data that might be relevant to solving a given problem must be known to the program. The key problem here is establishing a base of synonyms and knowledge about classes of data. To do this, we have been collecting protocols of students solving problems, requiring them to request all by simple initial case information. Student behavior also suggests disease knowledge that must be added to the knowledge base that an expert might not consider, but which the modeling program must recognize in a student. In general, we find that students carry out a much broader, inefficient search, requesting much more information than an expert and drawing fewer conclusions from the information that they receive.

The Image Student Modeling Program -- Teaching diagnosis involves recognizing the intent behind a student's behavior, so that missing knowledge can be distinguished from inappropriate strategies. The teacher ***interprets*** behavior, ***critiques*** it, and provides

advice about other approaches. To do this successfully and efficiently in a complex domain, the teacher benefits from multiple, complementary modeling strategies. IMAGE is a student modeling program that uses NEOMYCIN's meta-rules and disease knowledge to understand student diagnostic plans.

A student is presented with a problem to diagnose. As the student requests more problem data (i.e., takes a history and physical of the patient), IMAGE looks for regularities in sequences of his data requests. IMAGE contains a body of knowledge about how to map such sequences of behavior onto a strategical interpretation of what the student is doing. The process is heuristic in nature because the program will sometimes lose track of what the student is doing, because he is being inconsistent or using unexpected strategies.

The IMAGE uses a dual search strategy. The program first produces multiple predictions of student behavior by a model-driven simulation of NEOMYCIN. Focused, data-driven searches then explain incongruities. By supplementing each other, these methods lead to an efficient and robust plan understander.

A model of student strategies in medical diagnosis must disambiguate the possible purposes and knowledge underlying the student's actions. The approaches followed by other plan recognizers and student modelers are not sufficient here because:

1. the complex domain makes thorough searches impractical, whether top-down or bottom-up;
2. we are not modeling only facts and rules used in isolation, but also the procedures for applying them;
3. every one of the student's actions must be monitored in case the teaching module decides to interrupt;
4. his behavior must be evaluated and not just explained; and
5. we might not have any explicit goal statements from the student, so we expect to rely only on his queries for problem data as evidence for his thinking.

The IMAGE program is a prototype system which is now being extended. Specifically, a more useful system would examine its own interpretations and strive for coherence. We are designing such a such a system now, using the "blackboard model" for posting interpretations that may change over time. The levels of this blackboard are: 1) the student's data requests, 2) a classification of question type (e.g., triggered, follow-up, hypothesis-directed, general), 3) a strategic interpretation in terms of NEOMYCIN's diagnostic procedure (tasks and meta-rules). By incorporating a strategic level of interpretation, this program can be expected to make significant contributions to our understanding and use of the blackboard model of interpretation. The first version of this program will seek to explain student behavior in terms of deletion and reordering of procedural knowledge, plus simple variations of disease knowledge (e.g., false data/hypothesis relations). Study of student protocols is now suggesting what kinds of variations are common that we might easily identify automatically.

C.1.2 The NEOMYCIN Explanation System

The initial explanation system of NEOMYCIN, now completed, enables the user to answer WHY and HOW questions during a consultation. That is, when the program prompts the user for new data, the user may ask WHY the data is being requested or HOW some strategic task will be (or was) accomplished. Unlike MYCIN's explanation

system, upon which this kind of capability is patterned, explanations in NEOMYCIN are in terms of the diagnostic plan, not just specific associations between data and diagnoses. The program can provide abstract and concrete paraphrases of strategy rules (based on canned text). We have begun the next phase, which is to answer WHY questions by condensing the entire line of reasoning. The program will use models of the user's disease and strategic knowledge, plus general explanation heuristics, to select the task and focus information that is most likely to be of interest. Prototypic user models are now implemented. Heuristics have been designed and include: 1) mentioning the last task whose focus (or argument) changed in kind (e.g., from a disease hypothesis to a finding request); 2) never mentioning tasks that are merely iterating over a list of rules, findings, or hypotheses; and 3) only mentioning tasks with a rule as an argument to programmers.

Related to our explanation condensations is an effort to teach the strategic language of tasks to students. For example, we will have students annotate a NEOMYCIN typescript in terms of tasks and foci, to help them recognize good strategic behavior. This requires a common language of what the tasks are, e.g., "grouping" and "asking general question." Rather than just marking annotating tasks, we seek the ***principles*** by which the tasks could be consistently structured into primitives and auxiliary. These same principles could be used by the explanation system for choosing tasks to mention. Our current theory is that these primitive or "interesting" operations correspond to meta-rules that establish a new focus.

C.1.4 Graphics for Teaching

We are continuing make extensive use of graphics in our programs. For example, we are implementing a program that will mostly automatize the protocol collection process (though we are cautious about how menus will bias student behavior, even when lists are very long and full of irrelevant findings). As part of our series of instructional programs, GUIDON-WATCH is now being implemented as a graphic system for watching NEOMYCIN's reasoning. For example, we can highlight the hypotheses under consideration and show graphically how the program "looks up" its hierarchies before refining hypotheses. The design of GUIDON-ANNOTATE is also mostly complete. It will allow a student to mark up a typescript of NEOMYCIN's behavior using the same language of tasks the program uses when explaining its own behavior; iconic menus are very useful to avoid natural language difficulties (though it is clear that the student will sometimes need to "talk back").

C.2 Research in Progress

The following projects are active as of June 1983 (see also near-term plans listed in Section III.A):

1. augmenting NEOMYCIN's disease knowledge so we can fairly evaluate the program's focussing strategies and evaluate IMAGE;

2. developing capability to automatically produce summary explanations of NEOMYCIN's reasoning.

3. development of GUIDON-WATCH and GUIDON-ANNOTATE for teaching NEOMYCIN's knowledge to students.

4. developing new student modeling program based on the blackboard model.

D. Publications Since January 1983

1. Hasling, D., Clancey, W.J., Rennels, G.: ***Strategic explanations in Consultation.*** Int J Man-Machine Studies, in press.

2. Clancey, W.J.: ***The advantages of abstract control knowledge in expert system design.*** Proceedings of AAAI-83, pages 74-78.

3. Clancey, W.J.: ***Acquiring, representing, and evaluating a competence model of diagnosis.*** In Chi, Glaser, and Farr (Eds.), THE NATURE OF EXPERTISE. In preparation. HPP-84-2.

4. Clancey, W.J. and E. H. Shortliffe.:*READINGS IN MEDICAL ARTIFICIAL INTELLIGENCE: THE FIRST DECADE.* Reading: Addison-Wesley, in press.

5. Clancey, W.J.: ***Classification Problem Solving.*** HPP-84-7. Submitted to AAAI-84.

E. Funding Support

Contract Title: "Exploration of Tutoring and Problem-Solving Strategies"
Principal Investigator: Bruce G. Buchanan, Adjunct Prof. Computer Science
Associate Investigator: William J. Clancey, Research Assoc. Computer Science
Agency: Office of Naval Research and
Army Research Institute (joint)
ID number: N00014-79-C-0302
Term: March 1979 to March 1985
Total award: $683,892

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in GUIDON and NEOMYCIN has been shown by the medical and computer science communities. We are frequently asked to demonstrate these programs to Stanford visitors or at meetings in this country or abroad. GUIDON is available on the SUMEX 2020. Physicians have generally been enthusiastic about these programs' potential and what they reveal about current approaches to computer-based medical decision making.

Perhaps our most significant project to disseminate our research via SUMEX in the past year has been the completion of a book, "Readings in Medical Artificial Intelligence: The First Decade," edited by Dr. Clancey and Dr. Shortliffe. All of the significant SUMEX-AIM products of the past decade are described in this collection. Each chapter is preceded by a one-age historical introduction. In addition, opening and closing chapters by the editors survey issues in the field and the promise of the future. A complete index should make the book of considerable educational value. Preparation of this volume has been greatly aided by use of editing and formatting programs available on SUMEX-AIM. Royalties for the book, beyond production costs, will be used to sponsor an invited lecture at a major AI national conference, such as AAAI.

As mentioned earlier, a physician joined our group this year to help us develop the disease knowledge of the program (our first collaborator, Tim Beckett, MD, died of cancer in July 1983). This physician has found the convenience of accessing SUMEX from his

laboratory or at home to be extremely important for finding time to test NEOMYCIN and to communicate with us by electronic mail.

B. Sharing and Interaction with Other SUMEX-AIM Projects

GUIDON/NEOMYCIN retains strong contact with the ONCOCIN project, as both are siblings of the MYCIN parent. These projects regularly share programming expertise and continue to jointly maintain large utility modules developed for MYCIN. In addition, the central SUMEX development group acts as an important clearing house for solving problems and distributing new methods.

C. Critique of Resource Management

In the winter of 1984, the SUMEX staff efficiently and effectively shifted our operation away from the center of campus to a professional office building adjoining the medical center. The placement and installation of LISP workstations proceeded smoothly. After a year with Ed Pattermann as director of SUMEX, we can report that the stability and excellence of the resource we have come to expect has been completely maintained. Very important to us, the RAVEN laser printer installed at our new site not only provides excellent-quality output, but as a machine devoted to the Heuristic Programming Project has eliminated the delays we were experiencing a year ago.

With the shift to personal machines, we are continuing to experience a few difficulties. The greatest problem appears to be inadequately debugged software from XEROX. In particular, Interlisp-D relies heavily on network capabilities and must be compatible with several operating systems. This transition to new kinds of hardware and software can be expected to continue for several years. Therefore, we are extremely reliant upon the availability of experienced systems support. We believe that additional SUMEX staff is necessary to accommodate growing community needs.

III. RESEARCH PLANS

A. Project Goals and Plans

Research over the next year will continue on several fronts, leading to several prototype instructional programs by early 1985.

1. Continue to develop the knowledge base so the program can understand and anticipate any reasonable approach to the cases chosen for teaching.

2. Test student modeling program on these cases, collecting data for further development of the program, as well as exploring about the range of student approaches to diagnosis.

3. Extend the explanation system to do full summaries. Incorporate modeling capabilities that relate inquiries to a user model. Provide explanations tailored to this interpretation of the motivation behind the user's inquiry.

4. Integrate current display capabilities into running NEOMYCIN consultation to show how the space of diagnoses is explored and how diagnostic tasks are generated. Develop these capabilities to explore forms of graphic explanation useful in tutoring. (GUIDON-WATCH)

5. Extend student modeling system to include heuristics for generating tests that will confirm and extend the model. Improve the model to include analysis of patterns in model interpretations, including dependency-directed "backtracking" in the belief system and some capability to critique the modeling rules. Relate this to knowledge acquisition research.

6. Work closely with medical students to package NEOMYCIN capabilities in a "workstation" for learning medical diagnosis, determining what mix of student and program initiative is desirable.

B. Long term plans: the GUIDON2 Family of Instructional Programs

We sketch here our general conception of the research we plan for 1984-87, specifically the construction of instructional systems that use NEOMYCIN. Our ideas are strongly based on recent proposals by JS Brown, particularly his paper "Process versus Product --A perspective on tools for communal and informal electronic learning" and some related papers that he wrote in 1983. The plan is to implement at least three of these programs (here called GUIDON-WATCH, GUIDON-MANAGE, and GUIDON-ANNOTATE).

The key idea is that NEOMYCIN provides a ***language*** by which a program can converse with a student about strategies and knowledge organization for diagnosis. NEOMYCIN's tasks and structural terms provide the vocabulary or parts of speech; the meta-rules are the grammar of the diagnostic process. We will construct different graphic, reactive environments in which the student can observe, describe, compare, and improve diagnostic behavior of himself and others. There are many shared, underlying capabilities that will be constructed in parallel and improved over time.

Our approach is to delineate clearly different kinds of interactions that a student might have with a program concerning diagnostic strategies. Thus, each instructional system (but one) has a name of the form GUIDON-<student activity>, where the name specifies what the student is doing (e.g., watching, telling). The programs can be made arbitrarily complex by integrating coaches, student models, and explanation systems. We try here to separate out these capabilities, trying to get at the minimum interesting activities we might provide for a student.

GUIDON-WATCH The simplest system allows a student to watch NEOMYCIN solve a problem, perhaps one supplied by the student. Graphics display the evolving search space, that is, how tasks, as operators, affect the differential (Differential ---(Question X)---> Differential'). The student can step through slowly and replay the interaction. He can ask for prosaic explanations and summaries of what the program is doing. The program will also indicate its task and focus for each data request. This introduces the student to the idea that the diagnostic process has structure and follows a certain kind of logic.

GUIDON-MANAGE In this system the student solves a problem by telling NEOMYCIN what task to do at each step. Essentially, the student provides the strategy and the program supplies the tactics (meta-rules) and domain knowledge to carry out the strategy. The program will in general carry through tasks in a logical way, for example, proceeding to test a hypothesis completely, and not "breaking" on FINDOUT or APPLYRULES (two low-level tasks that mainly test domain knowledge and not strategy). The program will not pursue new hypotheses automatically. However, the student will always see what questions a task caused the program to request, as well as how the differential changes. This activity leads the student to observe the entailments of strategies, helping him become a better observer of his own behavior. Here he shows that he knows the structural vocabulary that makes a strategy appropriate.

GUIDON-ANNOTATE This system allows the student to annotate a NEOMYCIN typescript, indicating the task and focus associated with each data request. The program will indicate, upon request, where the student is incorrect and which annotations are different from NEOMYCIN's, but still reasonable interpretations. The student will be

able to choose these tasks from a menu of icons, either linearly or hierarchically displayed, as he prefers. (Again, NEOMYCIN will annotate its own solutions upon request and allow replaying.) This activity gets the student to think strategically by recognizing a good strategy. In this way, he learns to recognize how strategies affect the problem space.

GUIDON-APPRENTICE This is a variant of NEOMYCIN in which the program stops during a consultation and asks the student to propose the next data request(s). The student is asked to indicate the task and focus he has in mind, plus the differential he is operating upon. The program compares this proposal to what NEOMYCIN would do. In this activity we descend to the domain level and require the student to instantiate a strategy appropriately.

GUIDON-DEBUG Here the student is presented with a buggy version of NEOMYCIN and must debug it. He goes through the steps of annotating the buggy consultation session, indicating what questions are out of order or unnecessary, indicating what tasks are not being invoked properly, and then trying out his hypothesis on a "repaired" system. He is asked to predict what will be different, then allowed to observe what happens. This activity teaches the student to recognize how a diagnostic solution can be non-optimal, further emphasizing the value of good strategy. It also provides him with key meta-cognitive practice for criticizing and debugging problem behavior.

GUIDON-SOLVE This is the complete tutorial system. The student carries through diagnosis completely, while a plan recognizer attempts to track what he is doing and a coach interrupts to offer advice. Here annotation, comparison, debugging, and explanation are all integrated to illustrate to the student how his solution is non-optimal. For example, the student might be asked to annotate his solution after he is done; this will point out strategic gaps in his awareness and provide a basis for critique and improvement. A "curriculum" based on frequent student faults and important things to learn will drive the interaction. In this activity, the student is on his own. Faced with the proverbial "blank screen," he must exercise his diagnostic procedure from start to finish.

GUIDON-GAME Two or more students play this together on a single machine. They are given a case to solve together, and each student requests data in turn. All students receive the requested information. When a student is ready, he makes a diagnosis, indicated secretly to the program while the others are not watching. He then drops out of the questioning sequence. However, he can re-enter later, but of course will be penalized. Afterwards, score is based on the number of questions asked and use of good strategy. The coach will indicate to weak players what they could learn from strong players, encouraging them to discuss certain issues among themselves. Variation: one person solves while one or more competing students annotate the solution and show where it could be improved. Variation: one team introduces a bug into NEOMYCIN (and predicts the effect) and the other team finds it (as in SOPHIE). This activity will encourage students to share their experiences and talk to and learn from each other about the diagnostic process.

C. Requirements for Continued SUMEX Use

Although most of the GUIDON and NEOMYCIN work is shifting to Xerox Dolphins and Dandelions (D-machines), the DEC 2060 and 2020 continue to be key elements in our research plan. Our primary use of the 2060 will be to develop the NEOMYCIN consultation system, possibly by remote ARPANET access. Because of address space limitations, the consultation program can be combined with explanation or student modeling facilities, but not both, as is required for GUIDON2 programs. We continue to use the 2020 for demonstrating the original GUIDON program. As always, the 2060 will be essential for work at home, writing, and electronic mail.

D. Requirements for Additional Computing Resources

The D-machine's large address space is permitting development of the large program that complex computer-aided instruction requires. Graphics will enable us to develop new methods for presenting material to naive users. We also plan to use the D-machine as a reliable, constant "load-average" machine, for running experiments with physicians and students. The development of GUIDON2 on the D-machine will demonstrate the feasibility of running intelligent consultation or tutoring systems on small, affordable machines in physicians' offices, schools and other remote sites.

We currently have access to 1 1/2 DOLPHINs. We expect that 3 full time programmers will need access to two full machines. We are keeping logs so we can begin to understand patterns of activity and how these "personal" machines can be effectively shared.

E. Recommendations for Future Community and Resource Development

As we shift our development of systems to personal LISP machines, such as the DOLPHIN, it becomes more difficult to access these programs remotely for access from our homes (so that we may work conveniently during the evenings and weekends) and from remote sites for collaboration and demonstration. This problem will be partly ameliorated by "dial-up" (modem) access to these machines, but the use of bit-mapped displays requiring a high-bandwidth makes the phone lines inadequate for our purposes. Further technological development of networks, probably involving access over cables, will be necessary.

As computer resources become more distributed, the need for a central machine does not diminish. Programs and knowledge bases continue to be shared, requiring high-speed network connections among computers and file servers. SUMEX-AIM's role will shiftly slightly over the next few years to accommodate these needs, but its identity as a central resource will only change in kind, not importance. Moreover, sophisticated printing devices, such as the Xerox RAVEN, must necessarily be shared, again using a network. Maintenance of this network and its shared devices will become a key activity for the SUMEX staff. Thus, while computing resources will be provided by the "outboard engines" of personal machines, the community will remain intricately linked and dependent on common, but peripheral, resources.

From this perspective, future resource development should focus on improving the capabilities of networks, file servers, and attached devices to respond to individual requests. For example, it is now common for 10% of a user's time at a personal machine to be spent waiting for a file server or printer to process a request. Multi-processing becomes a necessity in such an environment, so a request can be honored, while the user returns to continue his programming or editing.

II.A.1.3. HPP Core AI Research

Heuristic Programming Project
Principal Investigator: Edward A. Feigenbaum
Co-Principal Investigator: Bruce G. Buchanan
Department of Computer Science
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A./B. Rationale and Medical Relevance

Medicine and the biological sciences are knowledge-intensive with an exponential rate of growth in relevant knowledge. This means that problem solving of all sorts is becoming increasingly complex in these disciplines. Further, most problems are symbolic in nature rather than amenable to mathematical formulation and numerical solution. Artificial Intelligence (AI) methods have been focused on medical and biological problems for over a decade with considerable success. This is because, of all the computing methods known, AI methods are the only ones that deal explicitly with symbolic information and problem solving and with knowledge that is heuristic (experiential) as well as factual.

One particularly fast-moving area of AI is expert systems. An expert system is one whose performance level rivals that of an human expert because it has extensive domain knowledge (currently usually derived from an human expert); it can reason about its knowledge to solve difficult problems in the domain; it can explain its line of reasoning much as an human expert can; and it is flexible enough to incorporate new knowledge without reprogramming. Expert Systems draw on the current stock of ideas in AI, for example, about representing and using knowledge. They are adequate for capturing problem-solving expertise for many bounded problem areas. Numerous high-performance, expert systems have resulted from this work in such diverse fields as analytical chemistry, medical diagnosis, cancer chemotherapy management, VLSI design, machine fault diagnosis, and molecular biology. Some of these programs rival human experts in solving problems in particular domains and some are being adapted for commercial use. Other projects have developed generalized software tools for representing and utilizing knowledge (e.g., EMYCIN, UNITS, AGE, MRS, GLISP) as well as comprehensive publications such as the three-volume ***Handbook of Artificial Intelligence*** and books summarizing lessons learned in the DENDRAL and MYCIN research projects.

But the current ideas fall short in many ways, necessitating extensive further basic research efforts. Our core research goals, as outlined in the next section, are to analyze the limitations of current techniques and to investigate the nature of methods for overcoming them. Long-term success of computer-based aids in medicine and biology depend on improving the programming methods available for representing and using domain knowledge. That knowledge is inherently complex -- it contains mixtures of symbolic and numeric facts and relations, many of them uncertain; it contains knowledge at different levels of abstraction and in seemingly inconsistent frameworks; and it links examples and exception clauses with rules of thumb as well as with theoretical principles. Current techniques have been successful only insofar as they severely limit this complexity. As the applications become more far-reaching, computer programs will have to deal more effectively with richer expressions and much more voluminous amounts of knowledge.

This report documents progress on the basic or core research activities within the

Heuristic Programming Project (HPP), funded in part under the SUMEX resource as well as by other federal and industrial sources. This work explores a broad range of basic research ideas in many application settings, all of which contributes in the long term to improved knowledge based systems in biomedicine.

C. Highlights of Research Progress

In the last year, we made progress on several major topics of research. The style of research that we believe is most productive at this stage of development of AI is the experimental style. Thus, within the HPP we build systems that implement our ideas for answering (or shedding some light on) fundamental questions; we experiment with those systems to determine the strengths and limits of the ideas; we redesign and test more; we attempt to generalize the ideas from the domain of implementation to other domains; and we publish details of the experiments. In order to carry out this style of research, then, we select specific problems to help focus the general questions. Many of these specific problem domains are medical or biological. In this way we believe the HPP has made substantial contributions to core research problems of interest not just to the AIM community but to AI in general.

Progress is reported below under each of the major topics of our work. Citations are to HPP technical reports listed in the publications section.

1. *Knowledge representation*: How can the knowledge necessary for complex problem solving be represented for its most effective use in automatic inference processes? Often, the knowledge obtained from experts is heuristic knowledge, gained from many years of experience. How can this knowledge, with its inherent vagueness and uncertainty, be represented and applied?

   Work on the logic-based MRS and the rule-based NEOMYCIN systems continues, attracting wide interest within the AI community. Numerous copies of MRS have been sent to collaborators elsewhere who are experimenting with it on the own machines. The book on rule-based expert systems by Buchanan & Shortliffe was completed in this year.

   [See HPP technical memos HPP-83-26, HPP-83-28, HPP-83-29, HPP-83-34, HPP-84-1]

2. *Advanced architectures and Control*: What kinds of software tools and system architectures can be constructed to make it easier to implement expert programs with increasing complexity and high performance? How can we design flexible control structures for powerful problem solving programs?

   A major effort in exploring and understanding the Blackboard architecture has been undertaken. A new pilot project using this architecture was started in the domain of protein chemistry (see description of Jardetzky & Buchanan pilot project). We have also begun investigating Blackboard systems as a way of organizing expert systems to exploit concurrency. Initial work has begun using the HASP/AGE systems as an application example.

   [See HPP technical memos HPP-83-30, HPP-83-33, HPP-83-38, HPP-83-43, HPP-83-44, HPP-84-4, HPP-84-6]

3. *Knowledge acquisition*: How is knowledge acquired most efficiently from human experts, from observed data, from experience, and from discovery? How can a program discover inconsistencies and incompleteness in its knowledge base? How can the knowledge base be augmented without perturbing the established knowledge base?

We have continued to make progress on two on-going projects for learning by experience and learning by analogy, and have initiated work on three new systems for acquiring knowledge. Those three are learning by watching, learning from text, and learning rules & meta-rules inductively. All three of the new systems use medical problems as their test-domains.

[Preliminary results have been published in HPP-83-27, HPP-83-36, HPP-84-2, HPP-84-8.]

4. *Knowledge utilization*: By what inference methods can many sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? How can knowledge be used intelligently, especially in systems with large knowledge bases, so that it is applied in an appropriate manner at the appropriate time?

   These issues are being explored in the development of MRS (Meta-Representation System) where one of the roles of meta-knowledge is to guide the effective use of lower level knowledge. They are also central in the studies of Blackboard control systems and their use in concurrent expert systems.

   [See HPP technical memos HPP-83-26, HPP-83-28, HPP-83-30, HPP-83-33, HPP-83-38, HPP-84-1, HPP-84-2, HPP-84-6]

5. *Software Tools*: How can specific programs that solve specific problems be generalized to more widely useful tools to aid in the development of other programs of the same class?

   We have continued the development of new software tools for expert system construction and the distribution of packages that are reliable enough and documented so that other laboratories can use them. These include the old rule-based EMYCIN system, MRS, and AGE.

   [See HPP technical memos HPP-83-26, HPP-83-28, HPP-83-29, HPP-83-33]

6. *Explanation and Tutoring*: How can the knowledge base and the line of reasoning used in solving a particular problem be explained to users? What constitutes a sufficient or an acceptable explanation for different classes of users? How can knowledge in a system be transferred effectively to students and trainees?

   The NEOMYCIN program has undergone preliminary comparison with medical students' protocols to understand the extent to which its medical concepts match those of the students. Analysis of experts' problem solving has also been done. NEOMYCIN's explanation capabilities have been improved. New work on student modelling has started in order to test NEOMYCIN in the context of tutoring.

   [See HPP technical memos HPP-83-41, HPP-83-42, HPP-84-2, HPP-84-7]

7. *Planning and Design*: What are reasonable and effective methods for planning and design? How can symbolic knowledge be coupled with numerical constraints? How are constraints propagated in design problems?

   The Palladio system for assisting in the design of VLSI circuits has been demonstrated and results presented in major publications and conferences.

   [See HPP technical memos HPP-83-31, HPP-83-39, HPP-83-45, HPP-83-46, HPP-83-47, HPP-84-3, HPP-84-5]

8. *Diagnosis*: How can we build a diagnostic system that reflects any of several

diagnostic strategies? How can we use knowledge at different levels of abstraction in the diagnostic process?

Research on using causal models in a medical decision support system (NESTOR) was largely completed and will be published in the coming year. A second medical diagnosis program that uses causal models of renal physiology (AI/MM) was also substantially completed and will be published soon. We are investigating the process of diagnosis in electronics as well as in medicine. The major thrust of this work has been integrating causal models about, and the structure of, a computer system or systems of the human body.

[See HPP technical reports: HPP-83-32, HPP-83-37, HPP-83-40, 84-7]

D. Relevant Publications

HPP-83-26 Michael R. Genesereth, *"MRS Casebook"*, May 1983.

HPP-83-27 Thomas D. Dietterich and Ryszard S. Michalski, *"Discovering Patterns in Sequences of Objects"*, May 1983.

HPP-83-28 Michael R. Genesereth, *"A Meta-level Representation System"*, May 1983.

HPP-83-29 M. Grinberg, *"MRS Installation Instructions"*, May 1983. This report available only to those who have purchased the software system MRS.

HPP-83-30 Barbara Hayes-Roth, *"The Blackboard Architecture: A General Framework for Problem Solving?"* May 1983.

HPP-83-31 Harold Brown, Christopher Tong, Gordon Foyster, *"Palladio: An Exploratory Environment for IC Design"*, June 1983.

HPP-83-32 John Kunz, E.A.Feigenbaum, Bruce G. Buchanan, E.H. Shortliffe, *"Comparison of Techniques of Computer-Assisted Decision Making in Medicine"*. Submitted for publication in the **Pure and Applied Biostructure**. World Press, Singapore (1983).

HPP-83-33 Nelleke Aiello, *"A Comparative Study of Control Strategies for Expert Systems: AGE Implementation of Three Variations of PUFF"*, June 1983.

HPP-83-34 Jock Mackinlay, *"Intelligent Presentation: The Generation Problem for User Interfaces"*, March 1983.

HPP-83-36 Russell Greiner and Michael R. Genesereth, *"What's New? A Semantic Definition of Novelty"*, June 1983.

HPP-83-37 Robert Joyce, *"Reasoning About Time-dependent Behavior in a System for Diagnosing Digital Hardware Faults"*, August 1983.

HPP-83-38 Barbara Hayes-Roth, *"The Blackboard Model of Control"*, June 1983.

HPP-83-39 Jerry Yan, Gordon Foyster, Harold Brown, *"An Expert System for Assigning Mask Levels to Interconnect in Integrated Circuits"*, October 1983.

HPP-83-40 Benoit Mulsant and David Servan-Schreiber, *"Knowledge Engineering: A Daily Activity on a Hospital Ward"*, October, 1983.

HPP-83-41 (working paper) Diane Warner Hasling, *"Strategic Explanations for a Diagnostic Consultation System"*, in AAAI Proceedings 1983 pp. 157-161.

HPP-83-42 Wm. J. Clancey, *"GUIDON"*, November 1983.

HPP-83-43 Narinder Singh, *"MARS: A Multiple Abstraction Rule-Based System"*, December 1983.

HPP-83-44 H.Penny Nii, *"Signal-to-Symbol Transformation: Reasoning in the HASP/SIAP Program"*, December 1983.

HPP-83-45 (working paper) Christopher Tong, *"A Framework for Circuit Design"*, December 1983.

HPP-83-46 (working paper) J.J. Finger, Michael Genesereth, *"RESIDUE - A Deductive Approach to Design"*, December 1983.

HPP-83-47 (working paper) J.J. Finger, Michael Genesereth, *"Planning to Gather Information"*, December 1983.

HPP-84-1 Michael R. Genesereth, *"Partial Programs"*, January 1984. (Replaces HPP-81-6)

HPP-84-2 (working paper) Wm. J. Clancey, *I"Acquiring, Representing, and Evaluating a Competence Model of Diagnostic Strategy"*, February 1984.

HPP-84-3 (working paper) Gordon Foyster, *"A Knowledge-Based Approach to Transistor Sizing"*, March 1984.

HPP-84-4 (working paper) Jock Mackinlay, Michael R. Genesereth, *"Implicit Language "*,March 1984.

HPP-84-5 Jeffrey Rosenschein, Michael R. Genesereth, *"Communication and Cooperation"*, March 1984.

HPP-84-6 D.E. Smith, Michael R. Genesereth, *"Controlling Recursive Inferences"*, March 1984.

HPP-84-7 William J. Clancey, *"Classification Problem Solving"*, March 1984.

HPP-84-8 (author), *"The Role of Abstractions in Understanding Analogy"*, April 1984.

E. Funding Support

We are pursuing a broad core research program on basic AI research issues with support from not only SUMEX but also DARPA, NASA, NSF, and ONR. SUMEX provides some salary support for staff and students involved in core research and invaluable computing support for most of these efforts. Additional salary support comes from the sources listed below.

Agency: National Library of Medicine; 5 P01 LM 03395
Project Title: Biomedical Knowledge Representation

Principal Investigator: Edward A. Feigenbaum
Amount: $95,424 (Direct Costs only)
Period Covered: 7/1/83 - 6/30/84

Agency: Defense Advanced Research Projects Agency; N00039-83-C-0136
Project Title: Heuristic Programming Project
Principal Investigators: Edward A. Feigenbaum and Bruce G. Buchanan
Amount: $3,354,493
Period Covered: 10/1/82 - 9/30/85

Agency: Defense Advanced Research Projects Agency; N00014-81-K-0303
Project Title: Intelligent Agents
Principal Investigator: Edward A. Feigenbaum
Award Amount: $484,652
Period Covered: 3/1/81 - 2/28/84
(the follow-on is merged with N00039-83-C-0136)

Agency: Defense Advanced Research Projects Agency/Martin Marietta;
(pending)
Project Title: Intelligent Task Automation
Principal Investigators: Michael R. Genesereth
Amount: $297,626
Period Covered: 10/1/83 - 2/28/85

Agency: Office of Naval Research; N00014-79-C-0302
Project Title: Recognizing and Articulating Diagnostic Skills
in an Intelligent Tutoring System
Principal Investigator: Bruce G. Buchanan
Award Amount: $1,110,447
Period Covered: 3/15/79 - 3/14/85

Agency: Office of Naval Research; N00014-80-C-0609
Project Title: Automatic Induction of Strategic Rules
Principal Investigator: Douglas B. Lenat
Award Amount: $108,000
Period Covered: 6/1/82 - 5/31/84

Agency: Office of Naval Research; N00014-81-K-0004
Project Title: Research on Introspective Systems
Principal Investigator: Michael R. Genesereth and Edward H. Shortliffe
Award Amount: $511,748
Period Covered: 1/1/84 - 12/31/86

Agency: NASA Goddard Space Flight Center; NAG 5-261
Project Title: Planning in Uncertain and Unforgiving Situations
Principal Investigators: Bruce G. Buchanan (and Thomas O. Binford)
Award Amount: $55,029
Period Covered: 9/1/83 - 8/31/84

Agency: NASA-AMES Research Center; NCC 2-220
Project Title: Research on Advanced Knowledge-based System Architectures
Principal Investigator: Edward A. Feigenbaum
Award Amount: $90,000
Period Covered: 1/1/84 - 11/30/84 (support level pending for future years)

Agency: NASA-AMES Research Center; NCC 2-274
Project Title: Research on Knowledge Representation
Principal Investigator: Bruce G. Buchanan
Award Amount: $50,000
Period Covered: 10/1/83 - 12/31/84 (support level pending for future years)

Agency: National Science Foundation; IST-83-12148
Project Title: Information Structure and Use in Knowledge-Based Expert Systems
Principal Investigator: Bruce G. Buchanan and Edward H. Shortliffe
Award Amount: $330,138
Period Covered: 3/15/84 - 2/28/87

Agency: IBM; IBM/Stanford Joint Study
Project Title: The Use of Design Models in the Diagnosis of Computer Hardware
Principal Investigator: Edward A. Feigenbaum
Award Amount: $660,000
Period Covered: 10/1/82 - 9/30/85

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

We rely on the central SUMEX facility as a focal point for all the research within the HPP, not only for much of our computing, but for communications and links to our many collaborators as well. As a common communications medium alone, it has significantly enhanced the nature of our work and the reach of our collaborations. As SUMEX and the HPP acquire a diversity of hardware, including LISP workstations machines and smaller personal computers, we rely more and more heavily on the SUMEX staff for integration of these new resources into the local network system. The staff has been extremely helpful and effective in dealing with the myriad of complex technical issues and leading us competently into this world of decentralized, diversified computing.

III. RESEARCH PLANS

A. Project Goals and Plans

The Core Research Project focuses on understanding the roles of knowledge in symbolic problem solving systems ¨ its representation in software and hardware, its use

for inference, and its acquisition. We are continuing to develop new tools for system builders and to improve old ones. The research crosses a number of application domains, as reflected in the subprojects discussed earler, but the main issues that we are addressing in this research are those fundamental to all aspects of AI. We believe this core research is broadening and deepening the groundwork for the design and construction of even more capable and effective biomedical systems.

As mentioned above, although our style of research is largely empirical, the questions we are addressing are fundamental. The three major research issues in AI have, since its beginning, been knowledge representation, control of inference (search), and learning. Within these topics, we will be asking the following kinds of questions and as our work progresses, we hope to leave behind several prototype systems that can be developed by others in the medical community.

1. Knowledge Representation -- How can we represent causal models and structural information? What are the relative benefits of logic-based, rule-based, and frame-based systems? How can we represent temporal relations and events so that reasoning over time is efficient?
2. Knowledge Acquisition -- How can an expert system acquire new knowledge without consuming substantial time from experts? Can we improve the knowledge engineering paradigm enough to make a difference? Can automatic learning programs be designed that will work across many disciplines? Will cooperative man-machine systems be able to open the communication channel between expert and expert system?
3. Knowledge Utilization -- By what inference methods can a variety of sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? What is the nature of strategy and control information?

Plans for the Coming Year -- Several systems have been developed in recent years to serve as vehicles for knowledge engineering and research on knowledge representation and its use. Knowledge acquisition (including machine learning) and advanced architectures for AI will be the two areas of most new activity in the coming year. Research on these topics obviously must draw on on-going work in representation and control.

In particular, we will focus on

- Inductive learning of MYCIN-like rules from case data in the domain of diagnosing disorders where the chief complaint is jaundice;
- Learning from experience in domains where the means for interpreting new data are largely contained in the emerging (and thus incomplete and not wholly correct) theory;
- Learning by watching a medical expert diagnose cases presented by NEOMYCIN;
- Investigating complex signal understanding systems for ways to exploit and represent concurrency with a view toward hardware and software architectures that may be capable of several orders of magnitude improvement in performance.

B. Justification and Requirements for SUMEX Use

Core research is essential to the vitality of a national resource for artificial

intelligence applications in biomedicine. It provides the new ideas and tools to address the limitations of existing experimental systems. We believe that the technical reports and programs produced as part of our continuing scientific efforts are received with interest by the AIM, and larger AI, research communities.

We require a stable source of computing cycles and substantial file space for the myriad of sub-projects that make up HPP/SUMEX core research. We anticipate no special needs beyond those in evidence this past year.

C. Computing Resources Outside of SUMEX-AIM

For some of the research reported here, we use Xerox-1100 series Lisp workstations, some of which were purchased by the NIH for SUMEX use. We have also purchased additional computing resources for the community with DARPA and HPP gift funds, including a VAX 11/780, a VAX 11/750, a Symbolics LM-2, 4 Symbolics 3600's, a Xerox Dorado, 2 Xerox Dandelions, and overflow cycles on the SCORE 2060. We expect to purchase additional Lisp workstations with similar funding over the next year and a half.

II.A.1.4. MOLGEN Project

MOLGEN - Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing, and Modification

Prof. E. Feigenbaum and Dr. P. Friedland
Department of Computer Science
Stanford University

Prof. Charles Yanofsky
Department of Biology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The MOLGEN project has focused on research into the applications of symbolic computation and inference to the field of molecular biology. This has taken the specific form of systems which provide assistance to the experimental scientist in various tasks, the most important of which have been the design of complex experiment plans and the analysis of nucleic acid sequences. We are now moving into a new phase of research in which we explore the methodologies scientists use to modify, extend, and test theories of genetic regulation, and then emulate that process within a computational system.

Theory or model formation is a fundamental part of scientific research. Scientists both use and form such models dynamically. They are used to predict results (and therefore to suggest experiments to test the model) and also to explain experimental results. Models are extended and revised both as a result of logical conclusions from existing premises and as a result of new experimental evidence.

Theory formation is a difficult cognitive task, and one in which there is substantial scope for intelligent computational assistance. Our research is toward building a system which can form theories to explain experimental evidence, can interact with a scientist to help to suggest experiments to discriminate among competing hypotheses, and can then revise and extend the growing model based upon the results of the experiments.

The MOLGEN project has continuing computer science goals of exploring issues of knowledge representation, problem-solving, discovery, and planning within a real and complex domain. The project operates in a framework of collaboration between the Heuristic Programming Project (HPP) in the Computer Science Department and various domain experts in the departments of Biochemistry, Medicine, and Biology. It draws from the experience of several other projects in the HPP which deal with applications of artificial intelligence to medicine, organic chemistry, and engineering.

B. Medical Relevance and Collaboration

The field of molecular biology is nearing the point where the results of current research will have immediate and important application to the pharmaceutical and chemical industries. Already, clinical testing has begun with synthetic interferon and human growth hormone produced by recombinant DNA technology. Governmental reports estimate that there are more than 200 new and established industrial firms already undertaking product development using these new genetic tools.

The programs being developed in the MOLGEN project have already proven useful and important to a considerable number of molecular biologists. Currently several dozen researchers in various laboratories at Stanford (Prof. Paul Berg's, Prof. Stanley Cohen's, Prof. Laurence Kedes', Prof. Douglas Brutlag's, Prof. Henry Kaplan's, and Prof. Douglas Wallace's) and over 400 others throughout the country have used MOLGEN programs over the SUMEX-AIM facility. We have exported some of our programs to users outside the range of our computer network (University of Geneva [Switzerland], Imperial Cancer Research Fund [England], and European Molecular Biology Institute [Heidelberg] are examples). The pioneering work on SUMEX has led to the establishment of a separate NIH-supported facility, BIONET to serve the academic molecular biology research community with MOLGEN-like software.

C. Highlights of Research Progress

C.1 Accomplishments

The current year has seen the completion of the previous grant's research on experiment design and debugging and the beginning of our new work on theory formation. The highlights of this work are summarized in several categories below.

C.1.1 Cloning Experiment Design

The cloning advisory system is now operational. It utilizes the following basic strategy or skeletal plan for the design of all experiments: First, isolate the piece of DNA you wish to clone, second, select a vector to carry the clone, third, insert the DNA into the vector, fourth, select a host for expression of the hybrid molecule, fifth, insert the hybrid into the host, and sixth, select for the protein or nucleic acid product that was the eventual goal of the cloning experiment. Following this skeletal plan, the cloning knowledge base contains information on DNA isolation methods, cloning vectors, insertion methods, hosts, host insertion methods, and selection methods.

This knowledge base has been tested on a wide range of cloning experiments in various laboratories. Dr. Rene' Bach finished work on the knowledge base by concentrating on two areas: vector selection and simulation of biological operations. He researched and described the criteria needed to make expert choices among several dozen different DNA cloning vectors, viewing that choice as being the "key" decision in the skeletal plan that would constrain and motivate the other decisions. He also did extensive work on describing the procedural knowledge necessary to accurately model the changes to DNA structures that take place during the course of a cloning experiment. This modeling serves to make decision-making during plan refinement more accurate and is also an important part of the experiment debugging system described below.

C.1.2 Experiment Debugging Research SPEX (the name given to the current version of our skeletal planning system) keeps complete records of all decisions made during the course of designing an experiment. These include strategic decisions as to which general planning heuristics to employ and which domain-specific skeletal plans to use, as well as tactical decisions made in the course of choosing specific operators to instantiate a plan step. In addition, SPEX keeps a dynamic model of the world state as assumed after the execution of each plan step. During the last year, Mr. Armin Hakin made use of this comprehensive information to extend the SPEX system to include experiment debugging facilities.

Experiment designs fail for one of three major reasons: a technical mistake in the laboratory (added too much salt, stopped a reaction too soon, etc.), a knowledge base mistake in technique selection (for example, the wrong enzyme was chosen for a cutting

step), or a strategic error--all of the steps work individually, but the design as a whole is in error. Our experiment debugging system has demonstrated an ability to cope well with errors of the first two types, and partially with errors of the final type.

The system works by first acquiring a description of the failed experiment and its goals from a scientist. This is done through a special experiment editing and description component that was added to the Unit System. The debugging system then queries the user to determine the skeletal plan that led to the creation of the particular experiment design; this step may involve the creation of a new skeletal plan (thereby serving as a useful aid to knowledge acquisition) or it may be that an existing skeletal plan will serve. If it is a new skeletal plan, then the system tries to find errors of the third type from above by utilizing some general skeletal plan design heuristics (e.g. making sure appropriate preconditions are established).

The system refines the skeletal plan given the goals and conditions of the experiment in question. It compares its choices with those actually selected by the scientist. When the debugging system's choices differ from those of the scientist, the system determines whether the difference indicated a fatal flaw in the scientist's plan or merely reflected different optimality criteria among nearly equal possibilities.

Finally, the system examines its model of what changes should occur in the laboratory environment during the course of the experiment. It informs the scientist when measurable changes should occur and asks him to compare those to actual changes. When a step is found whose "before" and "after" states do not correspond to predicted changes, then that step is pointed out as being suspect to a technical error of type 1 above.

C.1.3 Research in Theory Formation, Modification, and Testing

The first goal of our new work in scientific theory discovery was to extensively study an existing example of the process. Professor Charles Yanofsky's work in elucidating the structure and function of regulation in the trp operon of E. coli provided us with an excellent subject that spanned twelve years of research, dozens of collaborators, and almost one hundred research papers.

We have conducted extensive interviews with Professor Yanofsky and many of his former students and collaborators. We have examined most of the relevant research papers. We believe we now have a good understanding of the three major classes of knowledge that were important in the discovery of the theory of regulation in the trp operon: knowledge about the relevant biological objects, knowledge about the techniques used to elicit new information, and discovery heuristics used to build new models. The major stages in the discovery process have been mapped out, and work has begun on constructing a knowledge base that will represent the state of the world at the beginning of the trp operon research.

C.2 Research in Progress

The theory discovery project has two major goals over the next several months: first, to complete construction of a knowledge base that can be used to model and simulate the structure-function relationships relevant to genetic regulation, and second, to complete initial design of a computational architecture for theory extension, improvement, and discovery.

C.2.1 Building a Simulatable Model

The initial knowledge base will contain information relevant to genetic regulation in general and to the trp operon system in particular. The information will relate both to structure, i.e. the physical characteristics of the biological objects, and to function, i.e. the operational characteristics of the biological objects. In addition, the procedural knowledge needed to relate structure to function will play an important part in the knowledge base.

The goal is to have a knowledge base that can be used "actively" to simulate the result of various possible changes in the underlying regulatory model. For example, a common experimental method for studying a biological system is to introduce a mutation which destroys the functionality of some piece of the system. The regulatory knowledge base should be able to simulate and describe the results of such a "deletion mutation."

C.2.2 Design of Discovery System Architecture

In parallel with our work on knowledge base construction, we are designing an initial architecture for theory proposal, extension, and correction. In human scientists we have observed at least four major types of reasoning during the cognitive process. The first is data-driven reasoning when the major goal is to explain individual experimental results. The second is theory-driven reasoning which occurs when a partial theory or model drives its own extension. The third type of reasoning involves looking at closely related biological systems (e.g, noticing a similar behavior in the his operon system). The final type of reasoning relates to more distant analogies; thinking of DNA polymerase moving along a nucleotide sequence as similar to a railroad engine moving along a set of tracks. Our discovery system architecture will be able to embrace all of these reasoning types. A blackboard-style hybrid architecture is our initial guess, but much theoretical and experimental work needs to be done before we are satisfied with our architectural decisions.

D. Publications

1. Bach, R., Friedland, P., Brutlag, D. and Kedes, L.: *MAXIMIZE, a DNA sequencing strategy advisor*. Nucleic Acids Res. 10(1):295-304, January, 1982.

2. Bach, R., Friedland, P., and Iwasaki, Y.: *Intelligent computational assistance for experiment design*. Nucleic Acids Res. 12(1):11-29, January, 1984.

3. Brutlag, D., Clayton, J., Friedland, P. and Kedes, L.: *SEQ: A nucleotide sequence analysis and recombination system*. Nucleic Acids Res. 10(1):279-294, January, 1982.

4. Clayton, J. and Kedes, L.: *GEL, a DNA sequencing project management system*. Nucleic Acids Res. 10(1):305-321, January, 1982.

5. Feitelson, J. and Stefik, M.J.: *A case study of the reasoning in a genetics experiment*. Heuristic Programming Project Report HPP-77-18 (working paper), May, 1977.

6. Friedland, P.: *Knowledge-based experiment design in molecular genetics.* Proc. Sixth IJCAI, August, 1979, pp. 285-287.

7. Friedland P.: *Knowledge-based experiment design in molecular genetics.* Stanford Computer Science Report STAN-CS-79-760 (Ph.D. thesis), December, 1979.

8. Friedland, P., Kedes, L. and Brutlag D.: *MOLGEN--Applications of symbolic computation and artificial intelligence to molecular biology.* Proc. Battelle Conference on Genetic Engineering, April, 1981.

9. Friedland, P.: *Acquisition of procedural knowledge from domain experts.* Proc. Seventh IJCAI, August, 1981, pp. 856-861.

10. Friedland, P., Kedes, L., Brutlag, D., Iwasaki, Y. and Bach R.: *GENESIS, a knowledge-based genetic engineering simulation system for representation of genetic data and experiment planning.* Nucleic Acids Res. 10(1):323-340, January, 1982.

11. Friedland, P. and Iwasaki Y.: *The concept and implementation of skeletal plans.* (To appear in J. Artificial Intelligence, 19834

12. Friedland, P., Armstrong, P., and Kehler, T.: *The role of computers in biotechnology.* BIO\TECHNOLOGY 565-575, September, 1983.

13. Iwasaki, Y. and Friedland, P.: *SPEX: A second-generation experiment design system.* Proc. of Second National Conference on Artificial Intelligence, August, 1982, pp. 341-344.

14. Martin, N., Friedland, P., King, J. and Stefik M.J.: *Knowledge base management for experiment planning in molecular genetics.* Proc. Fifth IJCAI, August, 1977, pp. 882-887.

15. Meyers, S. and Friedland, P.: *Knowledge-based simulation of regulatory genetics in bacteriophage Lambda.* Nucleic Acids Res. 12(1):1-9, January, 1984.

16. Stefik, M. and Friedland, P.: *Machine inference for molecular genetics: Methods and applications.* Proc. of NCC, June, 1978.

17. Stefik, M.J. and Martin N.: *A review of knowledge based problem solving as a basis for a genetics experiment designing system.* Stanford Computer Science Report STAN-CS-77-596, March, 1977.

18. Stefik, M.: *Inferring DNA structures from segmentation data: A case study.* Artificial Intelligence 11:85-114, December, 1977.

19. Stefik, M.: *An examination of a frame-structured representation system.* Proc. Sixth IJCAI, August, 1979, pp. 844-852.

20. Stefik, M.: *Planning with constraints.* Stanford Computer Science Report STAN-CS-80-784 (Ph.D. thesis), March, 1980.

E. Funding Support

The MOLGEN grant is titled: MOLGEN: Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing, and Modification. It is NSF Grant MCS-8310236. Current Principal Investigators are Edward A. Feigenbaum Professor of Computer Science and Charles Yanofsky, Professor of Biology. MOLGEN is currently funded from 11/83 to 10/84 at $139,215 including indirect costs as the first year of a three year grant.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

SUMEX-AIM continues to provide the bulk of our computing resources. 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. 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 strongly approve of the network-oriented approach to a programming environment that SUMEX has begun to evolve into. The ability to utilize LISP workstations for intensive computing while still communicate with all of the other SUMEX resources has been very valuable to our work. We see a satisfactory mode of operation where most programming takes place on the workstations and most electronic communications, information sharing, and document preparation takes place within the mature TOPS-20 environment. The evolution of SUMEX has alleviated most of our previous problems with resource loading and file space. Our current workstations are not quite fast nor sophisticated enough, but we are encouraged by the progress that has been made.

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 benefited by interaction with other projects at yearly meetings and through exchange of working papers and ideas over the system.

The ability for instant communication with a large number of experts in this field has been a determining factor in the success of the MOLGEN project. It has made possible the near instantaneous dissemination of MOLGEN systems to a host of experimental users in laboratories across the country. The wide-ranging input from these users has greatly improved the general utility of our project.

We find it very difficult to find fault with any aspect of the SUMEX resource management. It has made it easy for us to expand our user group, to give demonstrations (through the 20/20 adjunct system as well as the LISP workstations), and to disseminate software to non-SUMEX users overseas.

III. RESEARCH PLANS

A. Project Goals And Plans

Our current work has the following major goals

1. Build a knowledge base that can be used for regulatory system simulation purposes. The knowledge base will represent the current model of an explanatory theory. We have already scoped the contents of this knowledge base and have begun construction.

2. Use the simulation knowledge base to explain observations that are indeed explainable without changes to the current model.

3. Begin to recognize when observations are "interesting" in that they contradict, dramatically confirm, are or unpredictable by the current model.

4. Build a mechanism for postulating extensions or corrections to the current theory: a constrained regulatory theory generator. Here are where the major AI architectural decisions will be made.

5. Build a mechanism for evaluating alternative theories.

6. Test this entire structure on the evolving trp regulatory system. Experiment with different knowledge bases to see how discovery is altered by the availability of new techniques.

7. Test the structure on several other areas of genetics.

B. Justification and Requirements for Continued SUMEX Use

The MOLGEN project depends heavily on the SUMEX facility. We have already developed several useful tools on the facility and are continuing research toward applying the methods of artificial intelligence to the field of molecular biology. The community of potential users is growing nearly exponentially as researchers from most of the biomedical-medical fields become interested in the technology of recombinant DNA. We believe the MOLGEN work is already important to this growing community and will continue to be important. The evidence for this is an already large list of pilot exo-MOLGEN users on SUMEX.

We support with great enthusiasm the acquisition of satellite computers for technology transfer and hope that the SUMEX staff continues to develop and support these systems. One of the oft-mentioned problems of artificial intelligence research is exactly the problem of taking prototypical systems and applying them to real problems. SUMEX gives the MOLGEN project a chance to conquer that problem and potentially supply scientific computing resources to a national audience of biomedical-medical research scientists.

II.A.1.5. ONCOCIN Project

ONCOCIN Project

Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The ONCOCIN Project is one of many Stanford research programs devoted to the development of knowledge-based expert systems for application to medicine and the allied sciences. The central issue in this work has been to develop a program that can provide advice similar in quality to that given by human experts, and to insure that the system is easy to use and acceptable to physicians. The work seeks to improve the interactive process, both for the developer of a knowledge-based system, and for the intended end user. In addition, we have emphasized clinical implementation of the developing tool so that we can ascertain the effectiveness of the program's interactive capabilities when it is used by physicians who are caring for patients and are uninvolved in the computer-based research activity.

B. Medical Relevance and Collaboration

The lessons learned in building prior production rule systems have allowed us to create a large oncology protocol management system much more rapidly than was the case when we started to build MYCIN. We introduced ONCOCIN for use by Stanford oncologists in May 1981. This would not have been possible without the active collaboration of Stanford oncologists who helped with the construction of the knowledge base and also kept project computer scientists aware of the psychological and logistical issues related to the operation of a busy outpatient clinic.

C. Highlights of Research Progress

C.1 Background and Overview of Accomplishments This Past Year

In the following list we have summarized the research and performance goals for the program, citing those which have been completely or partially accomplished and indicating those that have yet to be achieved:

1. to assist with identification of current protocols that may apply to a given patient [not yet undertaken; will not be relevant until more protocols than lymphomas have been encoded for routine use]

2. to assist with determining a patient's eligibility for a given protocol [not yet undertaken];

3. to provide detailed information on protocols in response to questions from clinic personnel [a query system has been developed and described in the medical computing literature; this initial system is designed for those building a protocol knowledge base; later versions will be used by physicians themselves];

4. to assist with chemotherapy dose selection and attenuation for a given patient [fully implemented and evaluated for patients under treatment for lymphoma; breast cancer protocols were recently implemented and released for use in the clinic; protocols for oat cell carcinoma complete, but not yet tested for release];

5. to provide reminders, at appropriate intervals, of follow-up tests and films required by the protocol in which a given patient is enrolled [fully implemented and evaluated for patients under treatment using ONCOCIN];

6. to reason about managing current patients in light of stored data from previous visits of (a) the individual patients [partially achieved, but much work remains; new funding has recently allowed us to undertake this task], or (b) the aggregate of all "similar" patients [not yet attempted].

7. to transfer the prototype system from its current research computer to a professional workstation that provides a model for cost-effective dissemination of clinical consultation systems [this is presently one of our major efforts];

8. to encode and implement for use by ONCOCIN the commonly used chemotherapy protocols from our oncology clinic [to facilitate this effort, a protocol acquisition system called OPAL is currently under development];

9. to develop a program to represent the therapy planning processes of expert clinicians in order to suggest treatment for patients whose special clinical situation precludes following the standard protocol [this effort was recently funded, and research has just commenced].

During the first year of this research (1979-1980), we developed a prototype of the ONCOCIN consultation system, drawing from programs and capabilities developed for the EMYCIN system-building project. During that year, we also undertook a detailed analysis of the day-to-day activities of the Stanford Oncology Clinic in order to determine how to introduce ONCOCIN with minimal disruption of an operation which is already running smoothly. We also spent much of our time in the first year giving careful consideration to the most appropriate mode of interaction with physicians in order to optimize the chances for ONCOCIN to become a useful and accepted tool in this specialized clinical environment.

The following year we completed the development of a special interface program that responds to commands from a customized keypad. We also encoded the rules for one more chemotherapy protocol (oat cell carcinoma of the lung) and updated the Hodgkin's Disease protocols when new versions were released late in 1980; these exercises demonstrated the generality and flexibility of the representation scheme we had devised. Software protocols were developed for achieving communication between the interface program and the reasoning program, and we coordinated the printing routines needed to produce hard copy flow sheets, patient summaries, and encounter sheets. Finally, lines were installed in the Stanford Oncology Day Care Center, and, beginning in May 1981, eight fellows in oncology began using the system three mornings per week for management of their patients enrolled in lymphoma chemotherapy protocols.

During our third year (1981 - 1982) the results of our early experience with physician users guided both our basic and applied work. We designed and began to collect data for three formal studies to evaluate the impact of ONCOCIN in the clinic. This latter task required special software development to generate special flow sheets and to maintain the records needed for the data analysis. Towards the end of 1982 we also began new research into a ***critiquing model*** for ONCOCIN that involves "hypothesis

assessment" rather than formal advice giving. Finally, in 1982 we began to develop a query system to allow system builders as well as end users to examine the growing complex knowledge base of the program.

Our fourth year (1982-1983) saw the departure of Carli Scott, a key figure in the initial design and implementation of ONCOCIN, the promotion of Miriam Bischoff to Chief Programmer, and the arrival of Christopher Lane as our second scientific programmer. At this time we began exploring the possibility of running ONCOCIN on a single-user professional workstation and experimented with different options for data-entry using a "mouse" pointing device. Christopher Lane has become our expert on the Xerox workstations that we are using, and most of the systems work and conversion effort described in Section C.2 below was designed or implemented by him. In addition, since ONCOCIN had grown to such a large program with many different facets, we spent much of our fourth year documenting the system. During that year we also modified the clinic system based upon feedback from the physician-users, made some modifications to the rules for Hodgkin's disease based upon changes to the protocols, and completed several evaluation studies.

ONCOCIN continues to be used routinely in the Stanford Oncology Clinic. Although it was originally made available only on three mornings per week, we have expanded the available time so that ONCOCIN may be used any time that the SUMEX 2020 computer on which it runs is not reserved for use by other research groups. The continued dependence on this time-shared computer, however, has prevented us from using ONCOCIN in in many clinical problem areas (other than the lymphomas where clinics are held three mornings per week, and breast cancer where clinic is held one day per week) because of our inability to assure the system's availability with reasonable response time at times other than the three mornings per week that SUMEX allows us to reserve the 2020. It is this latter point that has accounted for our decision not to spend a great deal of time developing new protocols to run on the 2020. Instead we have pressed our effort to adapt ONCOCIN to run on professional workstations (specifically the Xerox 1108 "Dandelion") which can eventually be dedicated to full time clinic use. We envision these workstations as the model for eventual dissemination of this kind of technology, and have been granted additional funding from DRR for three years to support workstation development along with knowledge base development so that we can add all of the protocols in use at the Stanford oncology clinic to ONCOCIN.

During the project's fifth year, three new full-time staff members, three students, and a new oncologist (Dr. Joel Bernstein) have joined our group. We are pleased that Dr. Robert Carlson, who was our Clinical Specialist for the past two years, has continued his affiliation with both Stanford and our research group. In August of 1983, Larry Fagan returned to Stanford after completing his M.D. degree. He has taken over the duties of the ONCOCIN Project Director along with becoming the Co-Director of the newly formed Medical Information Sciences Program. Dr. Fagan is in charge of coordinating the day-to-day efforts of our research. An additional programmer, Jay Ferguson, joined our group in the fall to assist with the effort required to transfer ONCOCIN from SUMEX to the 1108 workstation. A fourth programmer, Joan Differding, has joined our staff to work on our protocol acquisition effort. Samson Tu, a graduate student in Computer Science, John Williams, a medical student, and Mark Nakamura an undergraduate, are now working on ONCOCIN as well.

Funding from the NLM will continue to support the more basic research activities regarding biomedical knowledge representation, knowledge acquisition, therapy planning, and explanation as it relates to the ONCOCIN task domain. A grant from the NLM to study the therapy planning process was received, and this work (led by Dr. Fagan) has

commenced. This research is investigating how to represent the therapy planning strategies used to decide treatment for patients on the oat cell carcinoma protocol who run into serious problems requiring consultation with the protocol study chairman. Dr. Branimar Sikic, a faculty member from the Stanford University Department of Medicine, and the Study Chairman for the oat cell protocol, is collaborating on this project. A prototype system is being developed by John Williams.

In the following sections we will list our research goals and summarize recent research and development activities in greater detail.

C.2 Goal: To transfer the oncology prototype from the SUMEX research computer to a professional workstation

We have concentrated on five steps in the process of transferring the program to a professional workstation, each of which is discussed below. The transfer is from the SUMEX mainframe DEC-2020 running the INTERLISP-10 computer language to the XEROX 1108 scientific processor (called a "D-machine") running the INTERLISP-D programming language.

C.2.1 Development of a new physician interface for the graphics-oriented workstation

A major key to ONCOCIN's acceptance is the ability of the program to interface in a convenient fashion with the physician users. To reach this end we have designed a special computer graphics interface, called the Interviewer, that combines an exact replica of the familiar paper record with an advanced use of electronic pointing devices and electronic feedback.

During the last year we made major improvements to the D-Machine Interviewer program. The ONCOCIN Interviewer now has the ability to display historical information, to move back to older information not currently displayed on the computer screen, and an improved ability to select choices through multi-layered menus. Internally, it has been improved with a region based window system which increases both speed and flexibility. The region based window system, the register input devices, and the formatting language interface (that describes how forms should be presented on the display) have been generalized to be usable by other portions of the ONCOCIN project (notably the OPAL knowledge acquisition interface described below).

C.2.2 Development of new program to connect the physician interface to the reasoning portion of the program

The ONCOCIN system uses a special design that allows the Interviewer program and the reasoning section of the program to operate independently. In order to coordinate the activities of these two programs, a special communication program, called the Interactor was designed and built.

The Interactor program provides a message passing facility between two or more Interlisp-D processes (sub-programs that can run at the same time). The form of the messages are specified by the programmer. The system further allows messages to processes running on different machines via the computer network called the ETHERNET. This will allow moving components of a large program from one to several machines in a way invisible to the programs themselves. The Interactor also has the ability to find other Interactors on the local communication network.

C.2.3 Development of new programs to improve the efficiency and capabilities of ONCOCIN

In order to speed up both versions of the ONCOCIN system, we have written a simple rule and control block compiler for ONCOCIN that converts rules and control blocks into Interlisp programs, and then into compiled Interlisp. This helps to alleviate a memory space problem we have had in the Interlisp-10 version of the system as well as give us increased speed in the workstation version of the program.

Another systems level aspect of our work is in the creation and access of efficient patient record data files. To this end, we have implemented a machine independent hash file system (special data record format) that allows access to the data base via memory from disk files. The system is compatible with both Interlisp-10 and Interlisp-D and allows sharing of files between the two systems. Its format is also machine independent enough to allow access from other lisps on other computers. It is currently accepted by XEROX as a standard for the D machines and has been used by them to bring up programs of use by all D machine users. Along the same lines, we have experimented with solutions to the problems of having portions of text easily accessible by key from a file in a machine independent way.

C.2.4 Reorganization and recoding of existing programs for improved efficiency

The reasoning portion of the ONCOCIN program is being reprogrammed to increase speed and to benefit from the special capabilities of the Interlisp workstation. We are also re-writing parts of the program that were borrowed from other expert systems developed by our group.

We have reorganized the system into logical subsystems that are of a manageable size. This consisted of categorizing all the system functions (portions of the program) that are necessary for the Reasoner to run and putting each in an appropriate file. The Reasoner now runs in stand alone mode independently of which system it is on.

We are now in the process of cleaning up the specific programming part for each of the subsystems. This entails making various enhancements for both style and efficiency, adding comments and documentation, and further breaking down functionally independent parts of the system.

We have transferred portions of our EMYCIN utilities (based on the MYCIN expert system) and rewritten those utilities to make them work in both Interlisps. We have removed from the stand-alone Reasoner sections of the program that depended on the specific hardware of the DEC-2060 mainframe computer and now have versions of ONCOCIN on the 2060 and D-machines that are identical, being generated from the same program text. This step also included the use of the new hash file system (described above) on the D-machines.

C.2.5 System support for the reorganization

We have implemented a program called ***Graphcalls*** which allows programmers on the D-machines to visually graph the structure of the programs they have written. One can also examine the use of each of the functions on the graph as well as examine and change the variables they access. It also provides visual tracing and dynamic control of a program in execution. It has been used daily since its creation by both our project and members of the SUMEX community.

C.3 Goal: To modify the 2020 Clinic Version of ONCOCIN in response to user feedback

During the last year, we have added a number of new options to ONCOCIN for use by the fellows in the clinic. These include: a special option to request that a test be

ordered STAT (immediately), special menus for entering reasons for treatment modifications (these are used when there is a disagreement between the ONCOCIN recommended therapy and the physician's treatment plan, in order to gather data about why the physician has decided to override the system), and the option to request a copy of a patient's flowsheet be printed out on the clinic line-printer. We have also streamlined the methods that the various forms are created by ONCOCIN.

C.4 Goal: To encode and implement for use by ONCOCIN the commonly used chemotherapy protocols from our oncology clinic

We have pursued two approaches to increasing the number of protocols known by our system. The first approach is to use the existing software to implement active protocols not encoded at the time of our last report. The second approach has been to develop new software that is able to dramatically speed up the entry of protocols by providing graphically-oriented forms to be filled out on the computer that follow the basic outline of the protocol documents.

In the past, adding a new protocol to the ONCOCIN knowledge base has been a tedious process in which an oncologist and a programmer sit down and translate the oncologist's knowledge about the protocol into rules accessible to ONCOCIN. All the rules pertaining to the new protocol are written at that time, and this process must be repeated for every new protocol that is added to ONCOCIN. This method is rather inefficient since many of the rules are similar between protocols, differing only in their data content. To speed the process of knowledge acquisition, a program is being developed whereby a doctor could sit down at a terminal and fill in a series of forms containing appropriate questions about a new protocol. The information entered would take care of the large number of general rules pertaining to the protocol and allow the doctor and programmer to concentrate on the special cases.

The program will have two levels, the first of which is the program that will interact directly with the doctor. This program runs on Xerox D machines which have extensive graphics capabilities. Sections of the display screen (called windows) are organized in a way that emulates the physician's patterns of thought when thinking about the protocol. Other graphical entry devises have been used to encourage pointing at the answer rather than text entry. These methods are able to display all of the possible choices in a compact and comprehensible way. The first phase of this program has been completed and has been examined and approved by our oncology collaborators.

Information entered in the top level program will be converted to an intermediate data structure which will be used by the second level of the program to make new rules for the ONCOCIN knowledge base. Eventually, this process will also work in the opposite direction so that information about a previously entered protocol can be copied or modified by the physician for the new protocol. This "similar to" option will also extend to chemotherapies and drugs, so that when the doctor enters a chemotherapy or drug that the system knows about, pertinent information will be filled in for the doctor to copy or modify. The "similar to" capably along with the use of graphical input devices to speed the process of entering a new protocol and will also reduce errors and duplications. When this project is completed the total time needed to enter a new protocol should be greatly reduced and more effort will be concentrated on fine tuning the rules to handle special situations.

C.5 Evaluations of ONCOCIN's Performance

Data collection and analysis for all three ONCOCIN evaluations are now complete, results were presented at the annual meeting of the Society for Medical Decision Making, and we expect to have formal reports published during the next year.

Study 1, overseen by Dr. Robert Carlson of the Division of Oncology, is an evaluation of the program's impact on the attitude of the oncology fellows towards computers in general and ONCOCIN in particular. All physicians were administered questionnaires and structured interviews in the Spring of 1981 before ONCOCIN was introduced. The same questionnaires were distributed to them again after they had used the system for over a year. Follow-up interviews were also undertaken. This study was repeated again during 1983 to determine the trends over time. The results of this study are presently being prepared as a formal report.

We are also revising this study in preparation for the integration of the workstation version of ONCOCIN into the clinic. To maintain some consistency in the evaluation process, the original questions from Study 1 will be given and analyzed as before along with new questions. Several of these new survey and interview questions will serve as a "baseline" for evaluating any perceived improvements that will come with the introduction of the professional workstations in the clinic.

Study 2, overseen by Dr. Daniel Kent of the Division of General Internal Medicine, is an evaluation of the program's impact on the completeness and accuracy of flowsheet data recorded with and without ONCOCIN. Research programmers wrote routines to formally analyze on-line flow sheets for completeness and accuracy. Pre-ONCOCIN flow sheets were then entered into the system ***exactly as they were originally recorded by the physician***. The same analytic routines were used to analyze these pre-ONCOCIN flow sheets. The pre- and post- ONCOCIN data were compared. Results indicate that ONCOCIN has had a statistically significant beneficial impact on the completeness of data recording, the ordering of required tests, and the accuracy of the data recorded. A formal report of the results is in preparation.

Finally, Study 3 is examining the comparison between ONCOCIN's therapeutic advice and the treatment decisions made by oncology fellows in the same setting. The study was coordinated by Dr. David Hickam, formerly of our Division of General Internal Medicine and now on the faculty at the University of Oregon in Portland. Expert evaluators rated treatment plans without knowing whether the recommendation was that of ONCOCIN or one of the clinic physicians. Over 200 flow sheets were evaluated by Stanford lymphoma experts, and the resulting data have been fully analyzed by Dr. Hickam. The results indicate that the experts were unable to fault the recommendations made by ONCOCIN relative to those of experienced oncology fellows treating patients with lymphoma. A paper describing the results is in preparation.

A study was made of all of the cases run by physicians in the clinic to determine statistics about when they chose to override ONCOCIN's therapy recommendation. The results showed that approximately 75% of the time they agreed completely. When there were disagreements, 15% were about individual drug doses. This study pointed out a number of situations where ONCOCIN needs more knowledge, and where our expert needed clarification from the Principal Investigators of the particular protocol. As a result, a meeting was held (7/12/83) with some of the Faculty in charge of the Hodgkin's protocols to discuss issues arising from this study.

C.6 Documentation

An extensive effort to document the ONCOCIN system was completed during this last year. Many aspects of the ONCOCIN program and its programming environment are now written and available for project members' use. The increase in documentation has significantly reduced the start-up time for new researchers working with the project. In addition, we have published several papers and prepared several technical reports describing the system.

C.7 Hypothesis Assessment

As mentioned above, largely through the efforts of Curtis Langlotz, we have continued to develop modifications to ONCOCIN that will permit it to function as an "observer" of the physician's own decisions rather than as a primary source of advice. By permitting the physician to enter his or her own therapy plan on the flowsheet, we can acknowledge the oncologist's ability to reach appropriate therapeutic decisions for most patients. ONCOCIN will simply compare the physician's plan with what it believes is the proper therapy. If the system agrees with the physician, or determines that small differences are clinically insignificant, no advice from the computer will be necessary. If significant disagreements occur, on the other hand, ONCOCIN will need to respond with warnings and explanations for why it feels that an alternate therapy plan may be preferable. Our experience with ONCOCIN since its clinic implementation suggests that this mode of interaction will be preferred by the clinic physicians. It will require minimal changes to ONCOCIN's decision making approach, but the determination of what differences are clinically significant, and the optimal method for explaining their importance to the physician, are exciting challenges and important theoretical problems. An initial report describing this work appeared during 1983 in the ***International Journal of Man-Machine Studies***, and we plan to continue enhancing the system's critiquing and explanation capabilities. Mr. Langlotz presented this work in the 1983 Society for Computer Applications in Medical Care Conference Student Paper Competition, and was a finalist in the competition. The approach will not be used in the clinic, however, until ONCOCIN has been transferred to professional workstations, hopefully in about two years.

C.7 Query System and Rule Analysis

Shoko Tsuji has completed her work on the development of a query system to permit easy access to the large ONCOCIN knowledge base. Once we had encoded several hundred rules, it became unwieldy for system builders to work from large hard-copy listings of the knowledge base, and we anticipate that physicians will also require direct access to the program's knowledge. The query system permits this kind of access. Rather than dealing with natural language understanding by computer, we are designing ways that menu selection and the high-speed interface can be used to permit access to the information that is needed by a physician or system builder. A paper describing the early work was presented last year (May 1983) at the *AAMSI Congress 83* in San Francisco.

In previous reports we also described the work of Dr. Motoi Suwa who developed programs to assist in determining knowledge base consistency and completeness. His paper on this subject appeared in late 1982 in the *AI Magazine*. However, the programs that he wrote were never formally linked to our system for writing rules and modifying other parts of the knowledge base. As a result, Mr. Robert Noble spent time during the last few years modifying Suwa's code so that it would operate as an integral part of ONCOCIN. These changes have now been implemented so that a new rule can be dynamically compared to the rest of the knowledge base during the process of knowledge entry. Mr. Noble is currently considering how such a program might be implemented on a workstation in order to take advantage of the newly available graphical capabilities of these machines.

C.9 Encoding of Additional Protocols

As was indicated above, we have emphasized transfer of ONCOCIN to a professional workstation rather than the implementation of additional protocols. However, the oncologist in charge of breast cancer treatment at Stanford had expressed great interest in adding those treatment protocols to the system as soon as possible. We

have accordingly encoded and thoroughly tested the treatment plans for adjuvant therapy of breast carcinoma (CMF and CMFVP treatment plans) and released them for regular use in the spring of this year. Encoding of the CMF treatment plan required encoding of special rule types. In order to represent these treatment plans special methods were created for looking back to previous cycles to compare current laboratory results to previous values. This allows the development of treatment recommendations based upon past experience with the patient. A number of other protocols were added to the ONCOCIN system in order to keep the system's knowledge about Hodgkin's and Lymphoma protocols current. These included new Lymphoma protocols with very complex alternating chemotherapies (M-HOP/B-Cepp/HD-MTX and M-BACOD/HD-MTX), and new Hodgkin's protocols (alternating MOPP/ABVD).

C.10 Strategic Therapy Planning

As mentioned above, we have begun a new research project to study the therapy planning process, and how strategies which are used to plan therapy in difficult cases might be represented on a computer. This project, which we call the ONYX project, has as its goals: to conduct basic research into the possible representations of the therapy planning process; to develop a computer program to represent this process; and eventually to interface the planning program with ONCOCIN. The project members (Fagan, Bischoff, Williams, Langlotz, and Rennels) have spent many hours meeting with Dr. Sikic trying to understand how he plans therapy for patients whose special clinical situation precludes following the standard therapeutic plan described in the protocol document. In March of this year, the group spent two days at Xerox Palo Alto Research Center (PARC), working with Mark Stefik, Daniel Bobrow and Sanjay Mittal of PARC on possible representations for the knowledge structures and how such a program might run using the LOOPS knowledge programming system. We hope to have a prototype of this system running this year.

D. Publications Since January 1983

1. (*) Shortliffe, E.H. and Fagan, L.M. Expert systems research: modeling the medical decision making process. In: ***An Integrated Approach to Monitoring*** (J.S. Gravenstein, R.S. Newbower, A.K. Ream, and N.T. Smith, eds.), pp. 183-200, Woburn, MA: Butterworth's, 1983.

2. Duda, R.O. and Shortliffe, E.H. Expert systems research. ***Science,*** 220:261-268 (1983).

3. (*) Langlotz, C.P. and Shortliffe, E.H. Adapting a consultation system to critique user plans. ***International Journal of Man-Machine Studies,*** 19:479-496 (1983).

4. (*) Tsuji, S. and Shortliffe, E.H. Graphical access to the knowledge base of a medical consultation system. ***Proceedings of AAMSI Congress 83,*** pp 551-555, San Francisco, CA, May 1983.

5. (*) Bischoff, M.B., Shortliffe, E.H., Scott, A.C., Carlson, R.W. and Jacobs, C.D. Integration of a computer-based consultant into the clinical setting. ***Proceedings 7th Annual Symposium on Computer Applications in Medical Care***, pp. 149-152. October 1983, Baltimore, Maryland.

6. Mulsant, B. and Servan-Schreiber, D.: Knowledge engineering: a daily activity on a hospital ward. ***Computers and Biomedical Research*** 17:71-91 (1984).

7. (*) Shortliffe, E.H. Problems in implementing the computer for continuing education. Mobius. 3:52-55 (1983).

8. Shortliffe, E.H. The science of biomedical computing. In *Meeting the Challenge: Informatics and Medical Education* (J.C. Pages, A.H. Levy, F. Gremy, and J. Anderson, eds.), pp 1-10, Amsterdam, North-Holland, 1983.

9. Six Abstracts: Studies to Evaluate the ONCOCIN System.

- Shortliffe, E.H., Bischoff, M.B., Carlson, R.W., Jacobs, C.D. Clinical Integration to promote use and acceptance of a computer-based consultant. Presented at Annual Meeting Society for Medical Decision Making, Toronto, Canada, October 1983; reprinted in *Medical Decision Making* 3:358 (1983).

- Hickam, D.H., Shortliffe, E.H., Jacobs, C.D. A blinded evaluation of computer-based cancer chemotherapy treatment advice. *Clinical Research* 31(2):297A (1983).

- Hickam, D.H., Shortliffe, E.H. and Jacobs, C.D. An evaluation of the treatment recommendations of a computer-based cancer chemotherapy protocol advisor. Presented at Annual Meeting Society for Medical Decision Making, Toronto, 1983; reprinted in *Medical Decision Making* 3:362 (1983).

- Kent, D.L., Shortliffe, E.H., Bischoff, M.B. and Jacobs, C.D. The impact on quality of data management of a computer-based consultant program. Presented at Annual Meeting Society for Medical Decision Making, Toronto, October 1983; Reprinted in *Medical Decision Making* 3:362 (1983).

- Kent, D.L., Carlson, R.W., Jacobs, C.D. and Shortliffe, E.H. Evaluation of computer-based interactive data management for clinical trails. Presented at Annual meeting of the Western Section American Federation for Clinical Research, Carmel, February 1984; Reprinted in *Clinical Research* 32:31A (1984).

- Carlson, R.W., Shortliffe, E.H., Jacobs, C.D., Koretz, M.M. Physician attitudes toward a computer-based expert oncology consulting system. Submitted to Annual Meeting American Society for Clinical Oncology, Toronto, May 1984.

E. Funding Support

Grant Title: "Research Program: Biomedical Knowledge Representation"
Principal Investigator: Edward A. Feigenbaum
Co-Principal Investigator (ONCOCIN Project): Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: LM-03395
Term: July 1979 to June 1984
Total award: $497,420
Current award (1983-1984): $95,424

Grant Title: "Symbolic Computation Methods For Clinical Reasoning" (RCDA)
Principal Investigator: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: LM-00048
Term: July 1979 to June 1984

Total award: $196,425
Current award (1983-1984): $39,502

Grant Title: "The Development of Representation Methods to Facilitate Knowledge Acquisition and Exposition in Expert Systems"
Principal Investigator: Edward H. Shortliffe
Agency: Office of Naval Research
ID Number: NR 049-479
Term: January 1981 to December 1983
Total award: $456,622 (includes indirect costs)

Grant Title: "Studies in the Dissemination of Consultation Systems"
Principal Investigator: Edward H. Shortliffe
Agency: Biotechnology Resources Program, Division of Research Resources
ID Number: RR 01613
Term: July 1983 to June 1986
Total award: $624,455
Current award: (7/83-6/84): $220,371

Grant Title: "Therapy-planning strategies for consultation by computer"
Principal Investigator: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: LM-04136
Term: August 1983 to July 1986
Total award: $211,851
Current award: (8/83-7/84) $60,517

Grant Title: Henry J. Kaiser Faculty Scholar in General Internal Medicine
Principal Investigator: Edward H. Shortliffe
Agency: Henry J. Kaiser Family Foundation
Term: July 1983 to June 1986, renewable until June 1988
Total award: $150,000 ($50,000 annually).

Grant Title: Research on Introspective Systems
Principal Investigator: Michael R. Genesereth
Co-Principal Investigator: Edward H. Shortliffe
Agency: Office of Naval Research
ID Number: N00014-81-K-0004
Term: January 1, 1984 - December 31, 1986
Total award: $512,070 (includes indirects)

Grant Title: Information structure and use in knowledge-based expert systems
Principal Investigator: Bruce G. Buchanan
Co-Principal Investigator: Edward H. Shortliffe
Agency: National Science Foundation - IST83-12148
Term: March 1, 1984 - February 28, 1987
Total award: $330,000 (includes indirects)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in ONCOCIN has been shown by the medical, computer science, and lay communities. We are frequently asked to demonstrate the program to Stanford visitors (both the prototype system running in the clinic and the newer work transferring the system to professional workstations). We also demonstrated some of the developing workstation code on a machine loaned by the Xerox Corporation and installed at the *AAMSI Congress 83* held in San Francisco in May 1983. Physicians have generally been enthusiastic about ONCOCIN's potential. The interest of the lay community is reflected in the frequent requests for magazine interviews and television coverage of the work. Articles about MYCIN and ONCOCIN have appeared in such diverse publications as *Time* and *Fortune*, whereas ONCOCIN has been featured on the "NBC Nightly News", the PBS "Health Notes" series, and "The MacNeil-Lehrer Report." Due to the frequent requests for ONCOCIN demonstrations, we are producing a videotape about the ONCOCIN research which will include demonstrations of our the professional workstation research projects and the 2020-based clinic system. We have completed filming of the workstation demonstration programs and are ready to start filming the current clinic system and other necessary sequences. We expect the videotape to be complete by the summer and plan to make it available to interested SUMEX collaborators and other interested persons.

Our group also continues to oversee the MYCIN program (not an active research project since 1978) and the EMYCIN program. Both systems continue to be in demand as demonstrations of expert systems technology. MYCIN been demonstrated via networks at both national and international meetings within the last year, and several medical school and computer science teachers continue to use the program in their computer science or medical computing courses. We also have made the MYCIN program available to researchers around the world who access SUMEX using the GUEST account. EMYCIN has been made available to interested researchers developing expert systems who access SUMEX via the CONSULT account. One such consultation system for psychopharmacological treatment of depression, called Blue-Box, developed by two French medical students, Benoit Mulsant and David Servan-Schreiber, was reported on in July of 1983 in *Computers and Biomedical Research.*

B. Sharing and Interaction with Other SUMEX-AIM Projects

The community created on the SUMEX resource has other benefits that go beyond actual shared computing. Because we are able to experiment with other developing systems, such as INTERNIST/CADUCEUS, and because we frequently interact with other workers (at AIM Workshops or at other meetings), many of us have found the scientific exchange and stimulation to be heightened. Several of us have visited workers at other sites, sometimes for extended periods, in order to pursue further issues which have arisen through SUMEX- or Workshop-based interactions. In this regard, the ability to exchange messages with other workers, both on SUMEX and at other sites, has been crucial to rapid and efficient exchange of ideas. Certainly it is unusual for a small community of researchers with similar scholarly interests to have at their disposal such powerful and efficient communication mechanisms, even among those on opposite coasts of the country.

C. Critique of Resource Management

The transition from Tom Rindfleisch's able leadership to the directorship of Ed Pattermann went extremely smoothly, especially when one considers the simultaneous changeover to a new mainframe machine in early 1983. Our community of researchers has been extremely fortunate to work on a facility that has continued to maintain the high standards that we have praised in the past. The staff members are always helpful

and friendly, and work as hard to please the SUMEX community as to please themselves. As a result, the computer is as accessible and easy to use as they can make it. More importantly, it is a reliable and convenient research tool. We extend special thanks to Ed Pattermann for maintaining such high professional standards.

III. RESEARCH PLANS

A. Project Goals and Plans

In the coming year, there are seven areas in which we expect to expend our efforts on the ONCOCIN System:

1. We will complete preparation of, and submit for publication, papers describing the three ONCOCIN evaluations.
2. We will complete the filming and editing of a videotape about the ONCOCIN research.
3. We will continue to spend time maintaining the system's documentation and will prepare additional formal technical reports as well as the clinical reports on the evaluation studies.
4. We will continue to develop the hypothesis assessment approach to consultation (the ***critiquing model***) that was described above.
5. We will continue our efforts to transfer ONCOCIN to professional workstations and will begin planning for their implementation in the oncology clinic.
6. We will continue our efforts to develop a protocol acquisition system and begin to enter the other treatment protocols in use in the oncology clinic.
7. We will continue our basic research into the therapy planning process and develop a prototype system to assist with therapy planning.

B. Justification and Requirements for Continued SUMEX Use

All the work we are doing (ONCOCIN plus continued use of the original MYCIN program) is totally dependent on continued use of the SUMEX resource. Although some of the ONCOCIN work is shifting to Xerox workstations, the SUMEX 2060 and the 2020 continue to be key elements in our research plan. The programs all make assumptions regarding the computing environment in which they operate, and the ONCOCIN prototype in particular depends upon proximity to the DEC 2020 which enables us to use a 9600 baud interface.

In addition, we have long appreciated the benefits of GUEST and network access to the programs we are developing. SUMEX greatly enhances our ability to obtain feedback from interested physicians and computer scientists around the country. Network access has also permitted high quality formal demonstrations of our work both from around the United States and from sites abroad (e.g., Finland, Japan, Sweden, Switzerland).

We plan to continue development of ONCOCIN on both our own (recently purchased) Dandelion workstation and the shared SUMEX Dandelion workstation, and will be obtaining an additional workstation in the near future. However, the project now includes three graduate students (Langlotz, Tu, and Williams), two undergraduates (Noble, Nakamura), four full-time programmers (Bischoff, Differding, Ferguson, and Lane), a project director (Larry Fagan); in addition, new students will join us this summer and fall. Many of the students, and all of the programmers, need access to a workstation for major portions of their work. Due to the limited access to workstations, it will be necessary to continue use of the SUMEX 2060 for much of our work.

C. Requirements for Additional Computing Resources

The acquisition of the DEC 2020 by SUMEX was crucial to the growth of our research work. It has insured high quality demonstrations and has enabled us to develop a system (ONCOCIN) for real-world use in a clinical setting. As we have begun to develop systems that are potentially useful as stand-alone packages (i.e., an exportable ONCOCIN), the addition of personal workstations has provided particularly valuable new resources. We have made a commitment to the smaller Interlisp-D machines (Dandelions) produced by Xerox, and our work will increasingly transfer to them over the next several years. Our new funding will support our effort to implement ONCOCIN on workstations in the Stanford oncology clinic (and eventually to move the program to non-Stanford environments), but we will simultaneously continue to require access to Interlisp workstations made available by SUMEX for our research and development work. We are hopeful that it will be possible for SUMEX to commit to ONCOCIN considerable time on the new SUMEX workstations being acquired at the end of the current grant year.

The acquisition of the DEC 2060, coupled with our increasing use of workstations, has greatly helped with the problems in SUMEX response time that we had described in previous annual reports. We are extremely grateful for access both to the new central machine and to the research workstations on which we are currently building the new ONCOCIN prototype. The D-machine's address space is permitting development of the large knowledge base that ONCOCIN requires. The graphics capability of the workstations has also enabled us to develop new methods for presenting material to naive users. In addition, the D-machines have provided a reliable, constant "load-average" machine for running experiments with physicians and doing development work. The development of ONCOCIN on the Dandelion will demonstrate the feasibility of running intelligent consultation systems on small, affordable machines in physicians' offices and other remote sites.

D. Recommendations for Future Community and Resource Development

SUMEX is providing an excellent research environment and we are delighted with the help that SUMEX staff have provided implementing enhanced system features on the 2060 and on the workstations. We feel that we have a highly acceptable research environment in which to undertake our work. Workstation availability is becoming increasingly crucial to our research, and we have found over the past year that workstation access is at a premium. The SUMEX staff has been very helpful and understanding about our needs for workstation access, allowing us Dandelion use wherever possible, and providing us with systems-level support when needed. We look forward to the arrival of additional workstations and the development of a more distributed computing environment through SUMEX-AIM.

II.A.1.6. RADIX Project

The RADIX Project: Deriving Medical Knowledge from Time-Oriented Clinical Databases

Robert L. Blum, M.D., Ph.D.
Department of Computer Science
Stanford University

Gio C. M. Wiederhold, Ph.D.
Departments of Computer Science and Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Technical Goals - Introduction

Medical and Computer Science Goals -- The long range objectives of our project, called RADIX (formerly RX), are 1) to increase the validity of medical knowledge derived from large time-oriented databases containing routine, non-randomized clinical data, 2) to provide knowledgeable assistance to a research investigator in studying medical hypotheses on large databases, 3) to fully automate the process of hypothesis generation and exploratory confirmation. For system development we have used a subset of the ARAMIS database.

Computerized clinical databases and automated medical records systems have been under development throughout the world for at least a decade. Among the earliest of these endeavors was the ARAMIS Project, (American Rheumatism Association Medical Information System) under development since 1969 in the Stanford Department of Medicine. ARAMIS contains records of over 17,000 patients with a variety of rheumatologic diagnoses. Over 62,000 patient visits have been recorded, accounting for 50,000 patient-years of observation. The ARAMIS Project has now been generalized to include databases for many chronic diseases other than arthritis.

The fundamental objective of the ARAMIS Project as well as of all other clinical database researchers is to use the data that have been gathered by clinical observation in order to study the evolution and medical management of chronic diseases. Unfortunately, the process of reliably deriving knowledge has proven to be exceedingly difficult. Numerous problems arise stemming from the complexity of disease, therapy, and outcome definitions, from the complexity of causal relationships, from errors introduced by bias, and from frequently missing and outlying data. A major objective of the RADIX Project is to explore the utility of symbolic computational methods and knowledge-based techniques at solving some of these problems.

The RADIX computer program is designed to examine a time-oriented clinical database such as ARAMIS and to produce a set of (possibly) causal relationships. The algorithm exploits three properties of causal relationships: time precedence, correlation, and nonspuriousness. First, a Discovery Module uses lagged, nonparametric correlations to generate an ordered list of tentative relationships. Second, a Study Module uses a knowledge base (KB) of medicine and statistics to try to establish nonspuriousness by controlling for known confounders.

The principal innovations of RADIX are the Study Module and the KB. The Study Module takes a causal hypothesis obtained from the Discovery Module and produces a comprehensive study design, using knowledge from the KB. The study design is then executed by an on-line statistical package, and the results are automatically incorporated into the KB. Each new causal relationship is incorporated as a machine-readable record specifying its intensity, distribution across patients, functional form, clinical setting, validity, and evidence. In determining the confounders of a new hypothesis the Study Module uses previously "learned" causal relationships.

In creating a study design the Study Module follows accepted principles of epidemiological research. It determines study feasibility and study design: cross-sectional versus longitudinal. It uses the KB to determine the confounders of a given hypothesis, and it selects methods for controlling their influence: elimination of patient records, elimination of confounding time intervals, or statistical control. The Study Module then determines an appropriate statistical method, using knowledge stored as production rules. Most studies have used a longitudinal design involving a multiple regression model applied to individual patient records. Results across patients are combined using weights based on the precision of the estimated regression coefficient for each patient.

B. Medical Relevance and Collaboration

As a test bed for system development our focus of attention has been on the records of patients with systemic lupus erythematosus (SLE) contained in the Stanford portion of the ARAMIS Data Bank. SLE is a chronic rheumatologic disease with a broad spectrum of manifestations. Occasionally the disease can cause profound renal failure and lead to an early death. With many perplexing diagnostic and therapeutic dilemmas, it is a disease of considerable medical interest.

In the future we anticipate possible collaborations with other project users of the TOD System such as the National Stroke Data Bank, the Northern California Oncology Group, and the Stanford Divisions of Oncology and of Radiation Therapy.

We believe that this research project is broadly applicable to the entire gamut of chronic diseases that constitute the bulk of morbidity and mortality in the United States. Consider five major diagnostic categories responsible for approximately two thirds of the two million deaths per year in the United States: myocardial infarction, stroke, cancer, hypertension, and diabetes. Therapy for each of these diagnoses is fraught with controversy concerning the balance of benefits versus costs.

1. Myocardial Infarction: Indications for and efficacy of coronary artery bypass graft vs. medical management alone. Indications for long-term antiarrhythmics ... long-term anticoagulants. Benefits of cholesterol-lowering diets, exercise, etc.

2. Stroke: Efficacy of long-term anti-platelet agents, long-term anticoagulation. Indications for revascularization.

3. Cancer: Relative efficacy of radiation therapy, chemotherapy, surgical excision - singly or in combination. Optimal frequency of screening procedures. Prophylactic therapy.

4. Hypertension: Indications for therapy. Efficacy versus adverse effects of chronic antihypertensive drugs. Role of various diagnostic tests such as renal arteriography in work-up.

5. Diabetes: Influence of insulin administration on microvascular complications. Role of oral hypoglycemics.

Despite the expenditure of billions of dollars over recent years for randomized controlled trials (RCT's) designed to answer these and other questions, answers have been slow in coming. RCT's are expensive of funds and personnel. The therapeutic questions in clinical medicine are too numerous for each to be addressed by its own series of RCT's.

On the other hand, the data regularly gathered in patient records in the course of the normal performance of health care delivery are a rich and largely underutilized resource. The ease of accessibility and manipulation of these data afforded by computerized clinical databases holds out the possibility of a major new resource for acquiring knowledge on the evolution and therapy of chronic diseases.

The goal of the research that we are pursuing on SUMEX is to increase the reliability of knowledge derived from clinical data banks with the hope of providing a new tool for augmenting knowledge of diseases and therapies as a supplement to knowledge derived from formal prospective clinical trials. Furthermore, the incorporation of knowledge from both clinical data banks and other sources into a uniform knowledge base should increase the ease of access by individual clinicians to this knowledge and thereby facilitate both the practice of medicine as well as the investigation of human disease processes.

C. Highlights of Research Progress

C.1 1 May 1983 to 1 May 1984

Our primary accomplishments in this period have been the following:

1) complete modifications to RADIX to accommodate the one hundred-fold increase in the size of our database to 1700 patients,

2) carry out the study of the effect of prednisone on serum cholesterol on the new database,

3) publish results of the 1700 patient prednisone/cholesterol study,

4) publish the description of a two-stage regression method adapted by us to this study,

5) complete System Programmer's Manuals and User's Manual in preparation for transfer to outside sites, and

6) begin transfer of RADIX to Xerox D-Machine personal work stations.

C.1.1 Modifications to RADIX for the enlarged database

Extensive modifications to RADIX were required to deal with the 100-fold increase in the size of the database. The modifications necessary to run the study module automatically on the prednisone/cholesterol study were completed this year.

C.1.2 Prednisone/chlosterol study on enlarged database

We have carried out the automated study of the effect of prednisone on serum cholesterol using the new 1700 patient database. It has strongly confirmed the effect previously observed in the 50-patient SLE database. In addition, we are examining the effect in non-SLE patients and in other patient subsets. We are also examining alternative pharmacokinetic models for the prednione effect using the newly available data.

C.1.3 Publish results of prednisone/cholesterol study

The paper reporting these results is in draft form. It will be submitted for publication shortly.

C.1.4 Publish description of 2-stage regression method

A description of the 2-stage regression method has been submitted for publication.

C.1.5 Documentation

A two-volume System Programmer's Manual and a User's Manual describing implementation, maintenance and use of the system at Stanford has been completed. In addition, a complete set of the files needed for on-line demonstrations has been created, separating them from the working versions.

C.1.6 Transer of RADIX to D-Machines

Preliminary work on implementing RADIX on D-Machines has begun. This will continue in coming years.

C.1.7 Other accomplishments

We have presented the results of our research at several conferences during the year. Additional publications for the year are noted in the section on publications.

C.2 Research in Progress

We are currently completing additional studies on subsets of the 1700 patient database. These include automated analysis of the prednisone/ cholesterol effect in non-SLE patients and subsets of SLE patients, and fitting alternative pharmacokinetic models of the prednisone/cholesterol effect. This work should be completed shortly. We will then return to the more AI-oriented aspects of RADIX, as described below in the section on Research Plans.

D. Publications

1. Blum, R.L.: *Two Stage Regression: Application to a Time-Oriented Clinical Database.* (Submitted for publication to the American Journal of Epidemiology.)

2. Blum, R.L.: *Prednisone Elevates Cholesterol: An Automated Study of Longitudinal Clinical Data.* (Manuscript in preparation.)

3. Blum, R.L., and Walker, M.G.: *Minimycin: A Miniature Rule-Based System* (Submitted for publication to M.D.Computing)

4. Blum, R.L.: *Modeling and encoding clinical causal relationships.* Proceedings of SCAMC, Baltimore, MD, October, 1983.

5. Blum, R.L.: *Representation of empirically derived causal relationships.* IJCAI, Karlsruhe, West Germany, August, 1983 .

6. Blum, R.L.: *Machine representation of clinical causal relationships.* MEDINFO 83, Amsterdam, August, 1983.

7. Blum, R.L.: *Clinical decision making aboard the Starship Enterprise.* Chairman's paper, Session on Artificial Intelligence and Clinical Decision Making, AAMSI, San Francisco, May, 1983.

8. Blum, R.L. and Wiederhold, G.: *Studying hypotheses on a time-oriented*

database: An overview of the RX project. Proc. Sixth SCAMC, IEEE, Washington D.C., October, 1982.

9. Blum, R.L.: *Induction of causal relationships from a time-oriented clinical database: An overview of the RX project.* Proc. AAAI, Pittsburgh, August, 1982.

10. Blum, R.L.: *Automated induction of causal relationships from a time-oriented clinical database: The RX project.* Proc. AMIA San Francisco, 1982.

11. Blum, R.L.: *Discovery and Representation of Causal Relationships from a Large Time-oriented Clinical Database: The RX Project.* IN D.A.B. Lindberg and P.L. Reichertz (Eds.), LECTURE NOTES IN MEDICAL INFORMATICS, Springer-Verlag, 1982.

12. Blum, R.L.: *Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical database: The RX project.* Computers and Biomed. Res. 15(2):164-187, April, 1982.

13. Blum, R.L.: *Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project* (Ph.D. thesis). Computer Science and Biostatistics, Stanford University, 1982.

14. Blum, R.L.: *Displaying clinical data from a time-oriented database.* Computers in Biol. and Med. 11(4):197-210, 1981.

15. Blum, R.L.: *Automating the study of clinical hypotheses on a time-oriented database: The RX project.* Proc. MEDINFO 80, Tokyo, October, 1980, pp. 456-460. (Also STAN-CS-79-816)

16. Blum, R.L. and Wiederhold, G.: *Inferring knowledge from clinical data banks utilizing techniques from artificial intelligence.* Proc. Second SCAMC, IEEE, Washington, D.C., November, 1978.

17. Blum, R.L.: *The RX project: A medical consultation system integrating clinical data banking and artificial intelligence methodologies,* Stanford University Ph.D. thesis proposal, August, 1978.

18. Kuhn, Ingeborg, Gio Wiederhold, Jonathan E. Rodnick, Diane M. Ramsey-Klee, Sanford Benett, and Donald D. Beck: *Automated Ambulatory Medical Record Systems in the U.S.*, to be published by Springer-Verlag, 1983, in Information Systems for Patient Care, B. Blum (ed.), Section III, Chapter 14.

19. Walker, M.G., and Blum, R.L.: *A Lisp Tutorial.* (Submitted for publication to M.D.Computing.)

20. Wiederhold, Gio: *Knowledge and Database Management*, IEEE Software Premier Issue, Jan.1984, pp.63--73.

21. Wiederhold,Gio: *Networking of Data Information*, National Cancer Institute Workshop on the Role of Computers in Cancer Clinical Trials, National Institutes of Health, June 1983, pp.113-119.

22. Wiederhold, Gio: *Database Design* (in the Computer Science Series) McGraw-Hill Book Company, New York, NY, May 1977, 678 pp. Second edition, Jan. 1983, 768 pp.

23. Wiederhold, G.: IN D.A.B. Lindberg and P.L. Reichertz (Eds.), *Databases for Health Care*, Lecture Notes in Medical Informatics, Springer-Verlag, 1981.

24. Wiederhold, G.: *Database technology in health care.* J. Medical Systems 5(3):175-196, 1981.

E. Funding Support Status

1) Representation and Use of Causal Knowledge for Inference from Databases
Robert L. Blum, M.D., Ph.D.: Principal Investigator
Total award: $89,597 (direct + indirect)
Term: March 15, 1984 through March 14, 1986

2) Deriving Knowledge from Clinical Databases
Gio C. M. Wiederhold, Ph.D.: Principal Investigator
National Library of Medicine
Total award: $291,192 (direct)
Term: May 1, 1984 through November 30, 1986

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

During the last year we have completed System Programmer's Manuals and a User's Manual as steps towards making the system available to outside collaborators. We have had preliminary discussions with Drs. Edward Shortliffe and Lawrence Fagan concerning use of components of RADIX in the ONCOCIN project. Once the RADIX program is developed, we would anticipate collaboration with some of the ARAMIS project sites in the further development of a knowledge base pertaining to the chronic arthritides. The ARAMIS Project at the Stanford Center for Information Technology is used by a number of institutions around the country via commercial leased lines to store and process their data. These institutions include the University of California School of Medicine, San Francisco and Los Angeles; The Phoenix Arthritis Center, Phoenix; The University of Cincinnati School of Medicine; The University of Pittsburgh School of Medicine; Kansas University; and The University of Saskatchewan. All of the rheumatologists at these sites have closely collaborated with the development of ARAMIS, and their interest in and use of the RADIX project is anticipated. We hasten to mention that we do not expect SUMEX to support the active use of RADIX as an on-going service to this extensive network of arthritis centers, but we would like to be able to allow the national centers to participate in the development of the arthritis knowledge base and to test that knowledge base on their own clinical data banks.

B. Interactions with Other SUMEX-AIM Projects

Several of the concepts incorporated into the design of the RADIX Project have been inspired by other SUMEX-AIM Projects. The RADIX knowledge base is similar to the Units Package of the MOLGEN PROJECT. The production rule inference mechanism used by us is similar to that in the MYCIN Project.

Several programs developed by the MYCIN group are regularly used by RADIX. These include disk hash file facilities, text editing facilities, and miscellaneous LISP functions.

Regular communication on programming details is facilitated by the on-line mail system.

C. Critique of Resource Management

The DEC System 20 continues to provide acceptable performance, but it is frequently heavily loaded at peek hours.

The SUMEX resource management continues to be accessible and cooperative.

III. RESEARCH PLANS

A. Project Goals and Plans

The overall goal of the RADIX Project is to develop a computerized medical information system capable of accurately extracting medical knowledge pertaining to the therapy and evolution of chronic diseases from a database consisting of a collection of stored patient records.

SHORT-TERM GOALS --

For the last two years we have concentrated more heavily on publishing and presentation of our earlier AI results, on acquisition of a 1700 patient database, on medical studies based on the enlarged database, and on reporting the medical results and statistical techniques arising from our research. This is in concert with the long-term goal of ensuring that the work of the SUMEX / Artificial Intelligence in Medicine community be disseminated and applied in the general medical community.

During the coming two years we will concentrate much more on the artificial intelligence aspects of RADIX. We were successful this year in obtaining funding from the National Library of Medicine and the National Science Foundation to pursue this work. In particular, we will be deeply concerned with the representation of causal, temporal, and quantitative medical knowledge. It has become clear that these types of knowledge are crucial for the RADIX tasks of automated discovery of medical knowledge and the provision of intelligent automated assistance to clinical researchers, in addition to their generally perceived value in other medical expert systems applications.

LONG-RANGE GOALS -- There are two inter-related long-range goals of the RADIX Project: 1) automatic discovery of knowledge in a large time-oriented database and 2) provision of assistance to a clinician who is interested in testing a specific hypothesis. These tasks overlap to the extent that some of the algorithms used for discovery are also used in the process of testing an hypothesis.

We hope to make these algorithms sufficiently robust that they will work over a broad range of hypotheses and over a broad spectrum of data distributions in the patient records.

B. Justification and Requirements for Continued Use of SUMEX

Computerized clinical data banks possess great potential as tools for assessing the efficacy of new diagnostic and therapeutic modalities, for monitoring the quality of health care delivery, and for support of basic medical research. Because of this potential, many clinical data banks have recently been developed throughout the United States. However, once the initial problems of data acquisition, storage, and retrieval have been dealt with, there remains a set of complex problems inherent in the task of accurately inferring

medical knowledge from a collection of observations in patient records. These problems concern the complexity of disease and outcome definitions, the complexity of time relationships, potential biases in compared subsets, and missing and outlying data. The major problem of medical data banking is in the reliable inference of medical knowledge from primary observational data.

We see in the RADIX Project a method of solution to this problem through the utilization of knowledge engineering techniques from artificial intelligence. The RADIX Project, in providing this solution, will provide an important conceptual and technological link to a large community of medical research groups involved in the treatment and study of the chronic arthritides throughout the United States and Canada, who are presently using the ARAMIS Data Bank through the CIT facility via TELENET.

Beyond the arthritis centers which we have mentioned in this report, the TOD (Time-Oriented Data Base) User Group involves a broad range of university and community medical institutions involved in the treatment of cancer, stroke, cardiovascular disease, nephrologic disease, and others. Through the RADIX Project, the opportunity will be provided to foster national collaborations with these research groups and to provide a major arena in which to demonstrate the utility of artificial intelligence to clinical medicine.

C. Recommendations for Resource Development

The on-going acquisition of personal work-station Lisp processors is a very positive step, as these provide an excellent environment for program development, and can serve as a vehicle for providing programs to collaborators at other sites. Continued acquisitions are very desirable.

Another resource that would be highly desirable is a faster and more reliable means for transferring data and programs interactively between SUMEX and the CIT IBM 370. The addition of a reliable local network facility would greatly facilitate our ability to transfer patient files from CIT to SUMEX.

II.A.2. 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.

In addition to the progress reports presented here, abstracts for each project and its individual users are submitted on a separate Scientific Subproject Form.

II.A.2.1. CADUCEUS Project

CADUCEUS Project

J. D. Myers, M.D. and Harry E. Pople, Jr., Ph.D.
University of Pittsburgh
Decision Systems Laboratory
Pittsburgh, Pa., 15261

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

The principal objective of this project is the development of a high-level computer diagnostic program in the broad field of internal medicine as an aid in the solution of complex and complicated diagnostic problems. To be effective, the program must be capable of multiple diagnoses (related or independent) in a given patient.

A major achievement of this research undertaking has been the design of a program called INTERNIST-I, along with an extensive medical knowledge base. This program has been used over the past decade to analyze many hundreds of difficult diagnostic problems in the field of internal medicine. These problem cases have included cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in our Medical Center. In most instances, but by no means all, INTERNIST-I has performed at the level of the skilled internist, but the experience has high-lighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The institution of collaborative studies with other institutions has been deferred pending completion of the programs and knowledge base enhancements required for CADUCEUS. The installation of our own, dedicated VAX computer can be expected to aid considerably any future collaboration.

C. Highlights of Research Progress

----Accomplishments this past year

In a previous progress report the concept of "facets" of diseases was introduced and the need of CADUCEUS to proceed from broad pathophysio- logical and patho-biochemical concepts to specific disease processes was emphasized. The need for better representation of anatomical information and for better time representation were pointed out.

Drs. Miller and Myers have continued in the development of a new format for the CADUCEUS knowledge base. A major goal in making the transition from the INTERNIST-I knowledge base to that of CADUCEUS has been to insure that there is continuity between the two: The CADUCEUS knowledge base will be derived from the information in the INTERNIST-I knowledge base, with significant additions made as necessary.

A screen-oriented editor program for entering and manipulating the knowledge base was written in Franz Lisp. Using the editor, a total of 52 diagnosis nodes have been created and a total of 282 findings have been defined. Due to the more complex nature of a finding, the 282 findings represent over 600 old INTERNIST-I style manifestations.

In the CADUCEUS knowledge base, the basic unit of observational in- formation is called a finding. Unlike an INTERNIST-I manifestation, a finding can be assigned a status within a given patient, either "normal" or any number of forms of abnormal. For example, the status of the finding "heart murmur" can be either absent (normal) or present. Various qualifiers are allowed to modify a finding. For the finding "heart murmur", in a specific patient, a user might specify that it is heard at the second left interspace, that it is systolic, that it is heard in early systole only, that it is blowing, that it is grade 2 of 6, and its shape is crescendo- decrescendo. For findings whose values vary numerically, (e.g. SGOT-blood) the units of measurement and the normal range are specified so that a user may simply enter a number as the result of the test.

The concept of a disease profile has been carried over from the INTERNIST-I knowledge base. However, there are three separate diagnostic node types represented in the CADUCEUS knowledge base: the disease, the facet, and the subdivision. A disease is an entity whose presence should be reported if detected in a patient, and conceptually corresponds to the diseases mentioned in the separate chapters of standard medical textbooks. A subdivision is either a specific subtype of a disease (e.g. hepatitis B is a subtype of acute viral hepatitis, a disease) or a major specific organ system involvement by a multisystem disease (e.g. lupus nephritis and lupus cerebritis are subdivisions of the disease system lupus erythematosus).

The internal organization of disease, facet and subdivision profiles is identical. Apart from links to other nodes, there are nine essential components to each profile: disease parameters (e.g. prevalence of disease, specific sites it effects); demographic information about patients with the disease; general predisposing factors (which are interdependent, only one of which is likely to be present); independent risk factors (which often co-exist synergistically); general findings caused by the illness; specific findings which are relatively unique to the disease process; characteristic findings (e.g. a positive throat culture for beta hemolytic streptococcus in streptococcal pharyngitis); academically known but clinically contraindi- cated findings (e.g. one should not do a renal biopsy in patients with renal leptospirosis, but we know what the biopsy will show if it is done anyway); and manifestations whose presence make the diagnosis untenable (e.g. male sex makes pregnancy an invalid consideration). In addition to the aforementioned work in internal medicine, Drs. Gordon Banks and John Vries have been working on the development of a neurological diagnostic component for CADUCEUS. Dr. Banks has developed a neuroanatomic database which contains spatial descriptors for nearly 1,000 neuroanatomic structures and contains information as to their blood supply, and function. This database will allow anatomic localization of neurologic lesions. Some of this work for the peripheral nervous system has been done previously by students in our laboratory. The approach to the central nervous system has been to design a set of "symbolic coordinates". In constructing the neuroanatomic database, the human body, including the nervous system, is conceptually partitioned into a set of cubes (boxes). The largest cube, containing the entire body, is 2.187m on a side. This cube is divided into 27 smaller cubes, each 729mm on a side. Each of the smaller cubes is likewise subdivided until finally cubes that are each 1mm on a side are reached. Thus any cube has neighbors (of equal size) rostral, caudal, ventral, dorsal, left, and right of it, as well as a "parent" cube which contains it, and "daughter" cubes which it contains. Each of these cubes has the potential for being represented inside the computer program with a unique name (known as an atom in LISP, the language in which the database is programmed). Attached to each cube LISP atom

are lists of all of the anatomic structures that are completely and partially contained within the cube, as well as the blood supply to the region. This structure facilitates rapid retrieval of the location of a given anatomic structure as well as rapid localization of possible areas of involvement when there is evidence of dysfunction of one or more neural systems.

The hierarchical arrangement of the nested cubes ensures rapid convergence during searches, because if the sought object is not found in a parent cube, there is no need to search for it in any of the patient's children cubes. The addition of anatomic reasoning may allow parsimonious explanation of multiple manifestations arising from a single lesion, or allow the program to query the user regarding the presence of manifestations of involvement of areas that might be expected to be affected by whatever clinical state the program has under current consideration.

Dr. Vries has developed an imaging system using "octree encoding" to reconstruct n-dimensional images of the database as well and images of patients acquired by CT, NMR, and other neuroimaging techniques. Combining the database with the imaging system may open new areas of research, including clinical-pathological correlation of imaged lesions with symptoms, signs, and affected structures, automated reading of images, etc.

Dr. Miller in the last year completed work on a sub-project of CADUCEUS, called CPCS. He received support for this work from the National Library of Medicine New Investigator Program. The original objective of the project was to create a program, CPCS (for Computer-based Patient Case Simulator), to aid in the teaching of diagnosis to medical students. The INTERNIST-I/CADUCEUS knowledge base was to be used as the source of the program's medical expertise. This overall goal has been accomplished, and the program CPCS exists and runs on our VAX-11/780 using Franz Lisp. The CPCS project was a feasibility study to demonstrate that it is possible to construct a general case simulator. The project has been successful in that the CPCS program has been written, and runs quite well in its small test domain. But there is room for the future development of CPCS. Further construction of the CADUCEUS knowledge base, in areas beyond the current set of liver diseases, will significantly improve the utility of the CPCS program. As additional capabilities are added to CADUCEUS, the corresponding changes will be made in CPCS.

The medical knowledge base has continued to grow both in the incorporation of new diseases and the modification of diseases already profiled so as to include recent advances in medical knowledge. The knowledge base of 3/1/84 includes 591 individual disease profiles, 4,040 manifestations of disease, and about 3,500 "links" or interrelationships among diseases as well as a myriad of miscellaneous pieces of information which are essential for the correct operation of the system. Twenty new diseases have been profiled during the past year and the pediatrics knowledge base has continued to grow.

Recently the medical knowledge base (but not yet the diagnostic program) has been made available on line for use of the medical house staff at Presbyterian-University Hospital, our main teaching hospital in internal medicine, and at an affiliated community hospital in Pittsburgh, the Shadyside Hospital which operates a residency program in internal medicine. Preliminary reports indicate that the residents find the knowledge base useful.

----Research in progress

There are five major components to the continuation of this research project:

1. The enlargement, continued updating, refinement and testing of the extensive medical knowledge base required for the operation of INTERNIST-I.

2. The completion and implementation of the improved diagnostic consulting program, CADUCEUS, which has been designed to overcome certain performance problems identified during the past years of experience with the original INTERNIST-I program.

3. Institution of field trials of CADUCEUS on the clinical services in internal medicine at the Health Center of the University of Pittsburgh.

4. Expansion of the clinical field trials to other university health centers which have expressed interest in working with the system.

5. Adaptation of the diagnostic program and data base of CADUCEUS to subserve educational purposes and the evaluation of clinical performance and competence.

Current activity is devoted mainly to the first two of these, namely, the continued development of the medical knowledge base, and the implementation of the improved diagnostic consulting program.

D. List of relevant publications

1. Pople, Harry E.:*Knowledge-based Expert Systems: The Buy or Build Decision* IN Walter Reitman (Ed.), ARTIFICIAL INTELLI- GENCE APPLICATIONS FOR BUSINESS. Proceeding of the NYU Symposium. Ablex Pub. Corp., May 1983, pp. 23-40.

2. Myers, J.D.: *Artificial Intelligence and Medical Education* The Medical Journal, St. Joseph Hospital, Houston. Vol. 18, December 1983, pp. 193-202.

E. Funding support

1. Clinical Decision Systems Research Resource
Harry E. Pople, Jr., Ph.D.
Associate Professor of Business
Jack D. Myers, M.D.
University Professor (Medicine)
University of Pittsburgh
Division of Research Resources
National Institutes of Health

5 R24 RR01101-07
07/01/80 - 06/30/85 - $1,607,717
07/01/83 - 06/30/84 - $369,484

2. CADUCEUS: A Computer-Based Diagnostic Consultant
Harry E. Pople, Jr., Ph.D.
Associate Professor of Business
Jack D. Myers, M.D.
University Professor (Medicine)
University of Pittsburgh
National Library of Medicine
National Institutes of Health

5 R01 LM03710-04
07/01/80 - 06/30/85 - $817,884
07/01/83 - 06/30/84 - $196,710

3. Neurologic Consultation Computer Program
Gordon E. Banks, M.D.
Assistant Professor of Medicine
National Library of Medicine - New Investigator
National Institutes of Health

5 R23 LM03889-02
04/01/82 - 03/31/85 - $107,675
04/01/83 - 03/31/84 - $35,975
04/01/84 - 03/31/85 - $35,975

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A,B. Medical Collaborations and Program Dissemination Via SUMEX

CADUCEUS remains in a stage of research and development. As noted above, we are continuing 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.

C. Critique of Resource Management

SUMEX has been an excellent resource for the development of CADUCEUS. 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

A. Project Goals and Plans

Continued effort to complete the medical knowledge bases in internal medicine and pediatrics will be pursued including the incorporation of newly described diseases and new or altered medical information on "old" diseases. The latter two activities have proven to be more formidable than originally conceived. Profiles of added diseases plus other information is first incorporated into the medical knowledge base at SUMEX before being transferred into our newer information structures for CADUCEUS on the VAX. This sequence retains the operative capability of INTERNIST-I as a computerized "textbook of medicine" for educational purposes.

B. Justification and Requirements for Continued SUMEX Use

Our use of SUMEX will obviously decline with the installation of our VAX. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended for the present to keep INTERNIST-I at SUMEX for comparative use as CADUCEUS is developed here. Our team hopes to remain as a component of the SUMEX community and to share experiences and developments.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

Our predictable needs in this area will be met by the dedicated VAX computer recently installed.

D. Recommendations for Future Community and Resource Development

Whether a program like CADUCEUS, when mature, will be better operated from centralized, larger computers or from the developing self contained personal computers is difficult to predict. For the foreseeable future it would seem that centralized, advanced facilities like SUMEX will be important in further program development and refinement.

II.A.2.2. CLIPR - Hierarchical Models of Human Cognition

Hierarchical Models of Human Cognition (CLIPR Project)

Walter Kintsch and Peter G. Polson
University of Colorado
Boulder, Colorado

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The two CLIPR projects have made progress during the last year. The prose comprehension project has completed one major project, and is designing a prose comprehension model that reflects state-of-the-art knowledge from psychology (van Dijk & Kintsch, 1983) and artificial intelligence. During the last year, Polson, in collaboration with Dr. David Kieras of the University of Arizona have have continued work on a project studying the psychological factors underlying device complexity and the difficulties that nontechnically trained individuals have in learning to use devices like word processors. They have develop formal representations of a user's knowledge of how of operate a device and of the user-device interface (Kieras & Polson, in Press) and have completed several experiments evaluating their theory.

B. Technical Goals

The CLIPR project consists of two subprojects. The first, the text comprehension project, is headed by Walter Kintsch and is a continuation of work on understanding of connected discourse that has been underway in Kintsch's laboratory for several years. The second, the device complexity project is headed by Peter Polson in collaboration with David Kieras of the University of Arizona, Tucson. They are studying the learning and problem solving processes involved in the utilization of devices like word processors or complex computer controlled medical instruments (Kieras & Polson, in Press)

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by Kintsch and van Dijk (1978), although our programming efforts have identified necessary clarifications and modifications in that model (Miller & Kintsch, 1980, 1981; Kintsch & Miller, 1981; Miller, 1982). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension, and we are accordingly working with the AGE and UNITS groups at SUMEX toward the development of a knowledge-based, blackboard model of prose comprehension. We hope to be able to merge the substantial artificial intelligence research on these systems with psychological interpretations of prose comprehension, resulting in a computational model that is also psychologically respectable.

The goal of the device complexity project is to develop explicit models of the user-device interaction. They model the device as a nested automata and the user as a production system. These models make explicit kinds of knowledge that are required to operate different kinds of devices and the processing loads imposed by different implementations of a device. We feel that tools being developed at SUMEX--in particular AGE and the UNIT package--will dramatically facilitate our abilities to generate such models of the user-device interface.

C. Medical Relevance and Collaboration

The text comprehension project impacts indirectly on medicine, as the medical profession is no stranger to the problems of the information glut. By adding to the research on how computer systems might understand and summarize texts, and determining ways by which the readability of texts can be improved, medicine can only be helped by research on how people understand prose. Development of a more thorough understanding of the various processes responsible for different types of learning problems in children and the corresponding development of a successful remediation strategy would also be facilitated by an explicit theory of the normal comprehension process.

Note that our goal of a blackboard model is particularly relevant to the understanding of learning difficulties. One important aspect of a blackboard model is the separation of cognitive processes into a set of interacting subprocesses. Once such subprocesses have been identified and constructed, it would be instructive to observe the model's performance when certain of these processes are facilitated or inhibited. Many researchers have shown that there are a variety of cognitive deficits (insufficient short-term memory capacity, poor long-term memory retrieval, and such) that can lead to reading problems. Having a blackboard model in which the power of individual components could be manipulated would be a significant step in determining the nature of such reading problems.

The device complexity project has two primary goals: the development of a cognitive theory of user-device interaction in including learning and performance models, and the development of a theoretically driven design process that will optimize the relationships between device functionality and ease of learning and other performance factors (Polson & Kieras, 1983). The results of this project should be directly relevant to the design of complex, computer controlled medical equipment. We are currently using word processors to study user-device interactions, but principles underlying use of such devices should generalize to medical equipment.

Both the text comprehension project and the device complexity project involve the development of explicit models of complex cognitive processes; cognitive modelling is a stated goal of both SUMEX and research supported by NIMH.

The on-going development of the prose comprehension model would not be possible without our collaboration with the AGE and UNITS research groups. We look forward to a continued collaboration, with, we hope, mutually beneficial results. Several other psychologists have either used or shown an interest in using an early version of the prose comprehension model, including Alan Lesgold of SUMEX's SCP project, who is exporting the system to the LRDC Vax. We have also worked with James Greeno -- another member of the SCP project -- on a project that will integrate this model with models of problem solving developed by Greeno and others at the University of Pittsburgh. Needless to say, all of this interaction has been greatly facilitated by the local and network-wide communication systems supported by SUMEX. There has been considerable communication between members of the prose comprehension and AGE/UNITS groups as program bugs have been discovered and corrected; the presence of a mail system has made this process infinitely easier than if telephone or surface mail messages were required. The mail system, of course, has also enabled us to maintain professional contacts established at conferences and other meetings, and to share and discuss ideas with these contacts.

D. Progress Summary

The prose comprehension project has completed an initial version of a model of prose comprehension (Miller & Kintsch, 1980). This model has been applied to a large number of texts, and has yielded quite reasonable predictions of recall and readability. Psychologists from other universities have used this system to derive reading time and recall predictions for their own experimental materials. We are currently using the AGE and UNITS packages to extend this model toward one that can make use of world knowledge in its analyses; this model is discussed in Miller and Kintsch (1981) and Miller (1982). It is further developed in van Dijk and Kintsch (1983) has been applied to the domain of word arithmetic problems in our most recent work (Kintsch and Greeno, in Press).

The device complexity project is in it's third year. We have developed an explicit model for the knowledge structures involved in the user-device interaction, and we are developing simulation programs. Our preliminary theoretical results are described in Kieras & Polson (in Press). We have also completed several experiments evaluating the theory.

E. List of Relevant Publications

1. Kieras, D.E. and Polson, P.G.: ***An outline of a theory of the user complexity of devices and systems.*** Working Paper No. 1, Device Complexity Project, Universities of Arizona and Colorado, May, 1982.

2. Kieras, D.E. and Polson, P.G.: ***The formal analysis of user complexity.*** Int. J. Man-Machine Studies, In Press.

3. Kintsch, W. and van Dijk, T.A.: ***Toward a model of text comprehension and production.*** Psychological Rev. 85:363-394, 1978.

4. Kintsch, W. and Greeno, J.G.:***Understanding and solving word arithmetic problems.*** Psychological Review, In Press.

5. Miller, J.R. and Kintsch, W.: ***Readability and recall of short prose passages: A theoretical analysis.*** J. Experimental Psychology: Human Learning and Memory 6:335-354, 1980.

6. Miller, J.R. and Kintsch, W.: ***Readability and recall of short prose passages.*** Text 1:215-232, 1981.

7. Miller, J.R.: *A Knowledge-based Model of Prose Comprehension: Applications to Expository Text.* IN B.K. Britton and J.B. Black (Eds.), UNDERSTANDING EXPOSITORY TEXT. Erlbaum, Hillsdale, NJ, 1982.

8. Polson, P.G. and Kieras, D.E.: ***Theoretical foundations of a design process guide for the minimization of user complexity.*** Working Paper No. 3, Project on User Complexity, Universities of Arizona and Colorado, June, 1983.

9. Polson, P.G. and Kieras, D.E.: ***A formal description of users' knowledge of how to operate a device and user complexity.*** Behavior Research Methods and Instrumentation.

10. van Dijk, T.A. and Kintsch, W.:*STRATEGIES OF DISCOURSE COMPREHENSION.* Academic Press, New York, 1983.

F. Funding Support Status

1. Text Comprehension and Memory
   Walter Kintsch, Professor, University of Colorado
   National Institute of Mental Health - 5 Rol MH15872-14-16
   7/1/81 - 6/30/84: $281,085
   7/1/83 - 6/30/84: $69,878
2. Understand and solving word arithmetic problems
   Walter Kintsch, Professor, University of Colorado
   National Science Foundation
   8/1/83 - 7/31/86: $200,000

3. User Complexity of Devices and Systems
   David Kieras, Associate Professor, University of Arizona
   Peter G. Polson, Professor, University of Colorado
   International Business Machines Corporation
   1/1/82 - 12/31/84: $364,000
   1/1/84 - 12/31/84: $145,000

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Sharing and Interactions with Other SUMEX-AIM Projects

Our primary interaction with the SUMEX community has been the work of the prose comprehension group with the AGE and UNITS projects at SUMEX. Feigenbaum and Nii have visited Colorado, and one of us (Miller) attended the AGE workshop at SUMEX. Both of these meetings have been very valuable in increasing our understanding of how our problems might best be solved by the various systems available at SUMEX. We also hope that our experiments with the AGE and UNITS packages have been helpful to the development of those projects.

We should also mention theoretical and experimental insights that we have received from Alan Lesgold and other members of the SUMEX SCP project. The initial comprehension model (Miller & Kintsch, 1980) has been used by Dr. Lesgold and other researchers at the University of Pittsburgh, as well as researchers at Carnegie-Mellon University, the University of Manitoba, Rockefeller University, and the University of Victoria.

B. Critique of Resource Management

The SUMEX-AIM resource is clearly suitable for the current and future needs of our project. We have found the staff of SUMEX to be cooperative and effective in dealing with special requirements and in responding to our questions. The facilities for communication on the ARPANET have also facilitated collaborative work with investigators throughout the country.

III. RESEARCH PLANS

A. Long Range Projects Goals and Plans

The use of SUMEX by the prose comprehension group was greatly reduced in the

two years, because the focus of the work during that period was on experimental work and book writing, rather than computer simulation. This will change in the fall of 1984, when a new research associate will join the project whose primary responsibility will be in continuing the modelling work started in previous years with J. Miller (who is no longer associated with us). Thus, we expect a level of activity comparable to previous years next fall.

The primary goal of the device complexity project is the development of a theory of the processes and knowledge structures that are involved in the performance of routine cognitive skills making use of devices like word processors. We plan to model the user-device interaction by representing the users processes and knowledge as a production system and the device as a nested automata. We are also studying the role of mental models in learning how to use them.

B. Justification and Requirements for Continued SUMEX Use

The research of the prose comprehension project is clearly tied to continued access to the AGE and UNITS packages, which are simply not available elsewhere. We hope that our continued use of these systems will be offset by the input we have been and will continue to provide to those projects: our relationship has been symbiotic, and we look forward to its continuation.

C. Needs and Plans for Other Computational Resources

We currently use two other computing systems located at the University of Colorado. One is the Department of Psychology's VAX 11/780, which is used primarily to run real-time experiments to be modeled on SUMEX. The second is the University of Colorado's CDC 6400, which is used for various types of statistical analysis.

When the ARPA-sponsored Vax/Interlisp project is completed, we would be most interested in experimenting with becoming a remote AGE/UNITS site. It would seem that this sort of development is the ultimate goal of the package projects, and this type of interaction, once it becomes feasible, would be a logical extension of our association with the SUMEX facility.

D. Recommendations for Future Community and Resource Development

Our primary recommendation for future development within SUMEX involves (a) the continued support of INTERLISP, which is needed for AGE and for other work we have underway on SUMEX and (b) the continued development of the AGE and UNITS projects. In particular, we would like to see an extension of AGE to include a wider variety of control structures so that our psychological models would not be confined to one particular view of knowledge-based processing. The limited physical capacity of SUMEX, both in terms of address space and overloading, is, as before, a major problem. The prose comprehension group can no longer use the publicly released AGE/UNITS system due to its severely limited address space, and has had to build a personal AGE system from a stripped-down version of Interlisp and a selected subset of AGE and UNITS. We heartily endorse the plans underway to obtain more computing capacity for the SUMEX project.

Given our acquisition of a VAX, we particularly support the ongoing and continued development of INTERLISP for the VAX, so that local use of AGE and UNITS would be possible. Since we, as well as other psychologists, need the real-time capability of VAX/VMS to run on-line experiments, we hope that the INTERLISP system to be developed will be compatible with VMS. Note that this need for real-time work coincides with real-world applications of SUMEX programs, in which a VAX might be devoted to both real-time patient monitoring and diagnostic systems such as PUFF or MYCIN.

II.A.2.3. Rutgers Research Resource

Rutgers Research Resource--Computers in Biomedicine

Principal Investigators: Saul Amarel [1982-83], Casimir Kulikowski, Sholom Weiss [1983-84].
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 central 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.

At present, the total number of investigators who participate in scientific activities of the Resource is 83, of these, 20 have Rutgers appointments, 21 are outside investigators who participate in collaborative research projects that are mainly located at Rutgers, and 42 are investigators from collaborative national AIM projects that are located in different parts of the country. In addition, the Resource has 12 other members in Administrative, Computer Systems/Operations and general programming and secretarial functions. Thus, the Rutgers Resource community numbers at present a total of 95 participants.

Resource activities include research projects (collaborative research and core research) training/dissemination projects, and computing services in support of user projects.

B. Medical Relevance and Collaborations

In 1983-84 we continued the development of several versatile systems for building and testing consultation models in biomedicine. The EXPERT system has had many of its capabilities enhanced in the course of collaborative research in the areas of rheumatology, ophthalmology, and clinical pathology.

In ophthalmology we have developed a knowledge representation scheme for treatment planning which is both natural and efficient for encoding the strategies for choosing among competing and cooperating treatment plans. This involves a ranking of treatments according to their characteristics and desired effects as well as contraindications. Kastner has generalized the scheme so that it is now being used for a number of reasoning models: infectious eye disease, primary eye care, and rheumatology management. Our main collaboration continues to be with Dr. Chandler Dawson of the Proctor Foundation, UCSF.

In rheumatology, our collaboration with Drs. Donald Lindberg and Gordon Sharp at the University of Missouri-Columbia has continued at a very active level. The model for rheumatological diseases which now includes detailed diagnostic criteria for 26 major

diseases, had the management advice and treatment planning developed further. Dr. Sharp's group continues to develop the knowledge base in this area, with formalization of the knowledge carried out in conjunction with Dr. Lindberg's group and the Medical Expert Systems Group at Rutgers. The Resource researchers have developed new representational elements for EXPERT in response to the needs of the rheumatology research, and Politakis has developed a coordinated system called SEEK (System for Empirical Experimentation with Expert Knowledge) which provides interactive assistance to the human expert in testing, refining and updating a knowledge base against a data base of trial cases. SEEK has been tested and extended during the past year.

In clinical pathology our main collaboration has been with Dr. Robert Galen (Cleveland Clinic Foundation), with whom we have developed the serum protein electrophoresis model which is incorporated into an instrument - the scanning densitometer manufactured by Helena Laboratories. This instrument with interpretive reporting capabilities has now been on the market for over a year, is located at over 100 clinical sites, and represents the first known spin-off of AI expert systems research in the field of laboratory instrumentation. We continue to refine the representational mechanisms used for this kind of model.

In biomedical modeling applications we are experimenting with several prototype models for giving advice on the interpretation of experimental results in the field of enzyme kinetics, in conjunction with Dr. David Garfinkel. His PENNZYME program has been linked to a model in EXPERT, which allows the user to interpret the progress of the model analysis.

C. Highlights of Research Progress

Expert Medical Systems (C. Kulikowski, S. Weiss)

Research has continued on problems of representation, inference and control in expert systems. Emphasis has been placed this year on problems of knowledge base acquisition, empirical testing and refinement of reasoning (the SEEK system), and treatment planning strategies over time. From a technological point of view the market availability of the interpretive reporting version of a scanning densitometer, and the development of models for eye care consultation that run on microprocessor systems (Apple IIe, IBM-PC) represents an important achievement for AIM research in showing its practical impact in medical applications. This was recognized by the award of a scientific exhibit prize at the Academy of Ophthalmology Annual Meeting in November 1983.

1.1) SEEK: A System for Empirical Experimentation with Expert Knowledge

SEEK is a system which has been developed to give interactive advice about rule refinement during the design of an expert system. The advice takes the form of suggestions for possible experiments in generalizing and specializing rules in an expert model that has been specified based on reasoning rules cited by a human expert. Case experience, in the form of stored cases with known conclusions, is used to interactively guide the expert in refining the rules of a model. The design framework of SEEK consists of a tabular model for expressing expert-modeled rules and a general consultation system for applying a model to specific cases. This approach has proven particularly valuable in assisting the expert in domains where the logic for discriminating two diagnoses is difficult to specify; and we have benefited primarily from experience in building the consultation system in rheumatology.

1.2) Treatment Planning

The ranking and selection strategies developed as a stand-alone system last year

have been incorporated into the EXPERT framework. Capabilities for expressing reasoning over time have been added, so stored chart reviews can be carried out automatically, summarizing various patterns of findings over time, and abstracting the major features of interest for prognostic advice or treatment recommendations. Applications have been in infectious eye disease modeling, rheumatology treatment, and sequential advice in interpretation and sequencing of cardiac enzyme tests (e.g. CPK/LDH isoenzymes).

1.3) Technology Transfer

Important technology transfer milestones have also been achieved this year: the instrument interpretation EXPERT program for serum protein has been widely disseminated after being made available by Helena Laboratories, based on the prototype program developed by us; and we have succeeded in transferring a large knowledge base in rheumatology (about 1000 findings, 400 hypotheses and 1000 rules) onto a microprocessor (Motorola 68000) based system - the WICAT - which is well within the means of clinical researchers and practitioners. This system has been on site at the University of Missouri during the last year for testing and refining of the knowledge base.

1.4) Learning with Prior Structural Knowledge

This approach to knowledge acquisition and representation has as its goal to allow the expert to specify just the elements that are to enter into the reasoning model, with a few causal and taxonomic relations. These should then be sufficient to guide a learning program which operates on a data base of cases with known end-points. Such an approach would be useful in situations where the expert either has little time to explicitly formulate decision rules, or finds it difficult to do so. Our program [Drastal and Kulikowski, 1982] uses a blackboard representation, with multiple knowledge sources to handle the different conclusions, and the formation of rules from the data that pertain to them. We have tested this scheme in the areas of glaucoma and rheumatology, and shown that there are some interesting tradeoffs between the degree of a-priori structure provided by the expert, and the complexity of rule generation.

In relation to a system like SEEK, this approach represents a preprocessing or alternative means of developing the prototype model. We are now investigating the role of additional medical semantic constraints on the strategies of rule generation.

2) Artificial Intelligence:Expertise Acquisition and Problem Reformulation (S. Amarel)

The main research activity in this area is concerned with improvements in problem solving expertise via shifts in problem representation, i.e., via reformulation.

In this research, we have concentrated on the developmental processes that lead to the formation of specialized high performance procedures in sub-domains of a problem class. Theory formation is a key task in these processes; and we are now studying several approaches to this task - both top-down, model guided, approaches and 'bottom-up' methods that are based on detailed analysis of individual cases.

D. Up-to-Date List of Publications

The following is an update of publications in the Rutgers Resource for the period 1983 and 1984 (only publications not listed in previous SUMEX annual reports are presented here).

1. Weiss, S.M. and Kulikowski, C.A. *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

2. Kulikowski, C.A. contributor to the Knowledge Acquisition chapter edited by B. Buchanan in the book *Building Expert Systems* (F. Hayes- Roth, et al., eds) Addison-Wesley, 1983 (in press).

3. Yao, Y. and Kulikowski, C.A., *" Multiple Strategies of Reasoning for Expert Systems"*, Proc. Sixteenth Hawaii International Conference on Systems Sciences, pp. 510-514 , 1983.*

4. Kulikowski, C.A. *"Progress in Expert AI Medical Consultation Systems: 1980 - 1983 "*, Proc. MEDINFO '83 , pp. 499-502, Amsterdam, August 1983.*

5. Kastner, J.K., Weiss, S.M., and Kulikowski, C.A., *"An Efficient Scheme for Time-Dependent Consultation Systems"*, Proc. MEDINFO '83, pp.619-622, 1983.*

6. Kulikowski, C.A. *"Expert Medical Consultation Systems"*, Journal of Medical Systems, v.7, pp. 229-234, 1983.*

7. Weiss, S.M., Kulikowski, C.A., and Galen, R.S., *"Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program"*, Journal of Clinical Laboratory Automation, v.3, pp. 383-387, 1983.*

8. Kastner, J.K., Weiss, S.M., and Kulikowski, C.A., *" An Expert System for Front-line Health Workers in Primary Eye Care"*, Proc. Seventeenth Hawaii International Conference on Systems Sciences, pp. 162-166, 1984.*

9. Kulikowski, C.A. *"Knowledge Acquisition and Learning in EXPERT"*, Proc. 1983 Workshop on Machine Learning, Univ. of Illinois,Champaign-Urbana 1983.

Indicate by an asterisk (*) that the resource was given credit.

E. Funding Support

Since December 1983, the Rutgers Research Resource on Artificial Intelligence in Medicine is funded under grant RR 02230-01 from the Division of Research Resources, Biotechnology Resources Program. Principal Investigators are Casimir A. Kulikowski, Professor of Computer Science and Chairman of the Department of Computer Science [1984-87], and Dr. Sholom M. Weiss, Associate Research Professor of Computer Science.

The total direct costs for the period 1983-87 is $3,198,075, with the total for the current period (December 1, 1983 - November 30, 1984) being $ 989,276.

The Rutgers Resource was funded until December 1983 through an NIH grant entitled "Rutgers Research Resource on Computers in Biomedicine" - number P41RR643. The Co-Principal Investigators were Dr. Saul Amarel, Professor, Chairman of the Department of Computer Science, and Director of the Laboratory for Computer Science Research, and Dr. Casimir Kulikowski, Professor of Computer Science at Rutgers.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Dissemination

The SUMEX-AIM facility provides a backup node where some of our medical collaborators can access programs developed at Rutgers. The bulk of the medical collaborative work outlined in I.B. above is centered at the Rutgers facility (the Rutgers-AIM node).

Dissemination activities continue to be an important responsibility of the Rutgers Resource within the AIM community. The following activities took place in the last year:

1. Ninth AIM Workshop (1983):

   Organized by Dr. Casimir Kulikowski, it was held in Baltimore, in conjunction with the SCAMC 83 meeting. It consisted of a series of working group discussions followed by summary presentations by members of the AIM community on their conclusions.

2. Hawaii International Conference On Systems Sciences:

   Dr. Weiss presented a paper on the expert system for front-line health workers, and Dr. Kulikowski chaired a session on knowledge based medical systems.

3. VII-Pan-American Congress on Rheumatology:

   Dr. Sharp presented the rheumatology knowledge base and consultation program at this meeting.

4. At the AAAI-82 meeting, S. Amarel was elected member of the Executive Council of AAAI. He is also General Chairman of IJCAI-83 which was held in Karlsruhe, W. Germany in August 1983. Dr. Kulikowski was the organizer for an expert medical systems session at MEDINFO 83.

B. National AIM Projects at Rutgers

The national AIM projects, approved by the AIM Executive Committee, that are associated with the Rutgers-AIM node are the following:

1. INTERNIST/CADUCEUS project, headed by Dr. Myers and Dr. Pople from the University of Pittsburgh, has been using the Rutgers Resource as a backup system for development and experimentation.

2. Medical Knowledge Representation project, headed by Dr. Chandrasekaran from Ohio State University, is doing most of its research on the Rutgers system.

3. PURSUIT project, directed by Dr. Greenes from Harvard University, is doing most of its research on a Goal-Directed Model of Clinical Decision-Making at Rutgers.

4. Biomedical Modeling, by Dr. Garfinkel from the University of Pennsylvania.

5. Attending Project, directed by Dr. Perry Miller of the Yale Medical Center, is doing much of the research on critiquing a physician's plan of management at Rutgers.

6. MEDSIM project: This is a pilot project designed to provide resource-sharing and community building facilities for about 25 researchers in bio-mathematical modeling and simulation.

C. Critique of SUMEX-AIM Resource Management

Rutgers is currently using the SUMEX DEC-20 system primarily for communication with other researchers in the AIM community and with SUMEX staff, and

also for backup computing in demonstrations, conferences and site visits. Our usage is currently running at less than 50 connect hours per year at SUMEX, with an overall connect/CPU ratio of about 30.

Rutgers is beginning to place more emphasis on the use of personal computers, and on network support needed to make these effective. Sumex has been help in the following ways:

- The AIM Executive Committee allocated to the Rutgers-AIM node one of the Xerox Dolphins acquired by SUMEX, to help us develop experience in supporting personal machines. This machine was used almost entirely to help us develop and test network support(We are using Ethernet with the Xerox PUP networking protocols), and subsequently returned to SUMEX.
- Most of network software that we use was originally developed at SUMEX. Having this software available has saved us an enormous amount of time.
- Initially SUMEX was very helpful in giving us advice about setting up our Ethernet and the Dolphins.

III. RESEARCH PLANS

A. Project Goals and Plans

We are planning to continue along the main lines of research that we have established in the Resource to date. Our medical collaborations will continue with emphasis on development of expert consultation systems in rheumatology, ophthalmology and clinical pathology. The basic AI issues of representation, inference and planning will continue to receive attention. Our core work will continue with emphasis on further development of the EXPERT framework and also on AI studies in representations and problems of knowledge and expertise acquisition. We propose to work on a number of technology transfer experiments to micro processing that will be affordable by our biomedical research and clinical collaborators. We also plan to continue our participation in AIM dissemination and training activities as well as our contribution -- via the RUTGERS/LCSR computer -- to the shared computing facilities of the national AIM network.

B. Justification and Requirements for Continued SUMEX Use

Continued access to SUMEX is needed for:

1. Backup for demos, etc.
2. Programs developed to serve the National AIM Community should be runnable on both facilities.
3. There should be joint development activities between the staffs at Rutgers and SUMEX in order to ensure portability, share the load, and provide a wider variety of inputs for developments.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

Our computing is going to move in the direction of personal computers. We will continue to use Sumex for backup purposes, however.

D. Recommendations for Future Community and Resource Development

Use of personal computers and minicomputers is continuing to grow in the AIM community. We find that the biggest challenge is supporting these systems. Although some central computing will continue to be needed for communication and coordination, we believe that over the next few years all AIM research projects and even individual collaborators will come to have their own hardware. However many of these community members (particularly the collaborators) will not be in a position to support hardware or software on their own. We would certainly expect SUMEX to continue to provide expert advice in this area. However we believe it would be helpful for SUMEX to have a formal program to support smaller computers in the field. We envision this as including at least the following items:

- A central source of information on hardware and software that is likely to be of interest to the AIM community. SUMEX might want to become a distribution point for certain of this software, and even help coordinate quantity purchase of hardware if this proves useful.
- Assistance in support of hardware and software in the field. Depending upon the hardware involved, this might involve advice over the telephone or actual board-swapping by mail. With our Dolphins we have found that there are a number of problems that can be resolved over the telephone if we can find someone with appropriate expertise.

II.A.2.4. SECS: Simulation & Evaluation of Chemical Synthesis

SECS - Simulation and Evaluation of Chemical Synthesis Project

Principal Investigator: W. Todd Wipke
Board of Studies in Chemistry
University of California
Santa Cruz, CA. 95064

Coworkers:

I. Kim	(Grad student)
D. Rogers	(Grad Student)
J. Chou	(Postdoctoral)
M. Hahn	(Grad Student)
M. Yanaka	(Postdoctoral)
I. Iwataki	(Postdoctoral)

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

With the SECS project our long range goal 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 second area of research, the XENO project, is aimed at improving methods for predicting potential biological activity of metabolites and plausibility of incorporation and excretion of metabolites.

B. Medical Relevance and Collaboration

The development of new drugs and the study of drug structure biological activity relationships 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 bio-molecular 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 5) capability of computer to see molecules in a graph theoretical sense, free from the bias of 2-D projection.

The objective of using XENO in metabolism studies 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, one may envision applications of this technology in problem domains where one wishes to alter molecules in order to inhibit certain types of metabolism.

C. Highlights of Research Progress

C.1 SECS Project Developments

The majority of our research has been aimed at strategic planning in chemical synthesis. Specific work has included the SST project for recognizing potential starting materials from a target, the MCS project for maximal common subgraph searching, and a project for rapid substructure search using parallelism.

C.1.a SST -- Starting Material Strategies. The importance of selecting good starting materials for a synthesis has been known for a long time, but only recently has work started on applying computer techniques to the selection process. The selection of starting material for a synthesis is frequently the major discovery in a synthesis and the process of converting the starting material to the target is minor by comparison. Last year we reported development of the SST program for selecting starting materials that are appropriate for a given synthetic target using a library of available chemicals, but without reference to reactions. SST handles problems of classes I-III given below:

I) Target = SM	Identical match
II) Target > SM	Superstructure match
III) Target < SM	Substructure match
IV) None of these	Similarity match

For a search over our abstracted file, the identical match means that the target and starting materials are identical except for functionalization. The superstructure match is the case where we must ***make*** carbon-carbon bonds during a synthesis. The substructure match is the case where the starting material is larger than the target, so carbon-carbon bonds have to be ***broken***. Finally, the similarity match is where carbon-carbon bonds have to be both ***made and broken*** during the synthesis.

Our research in efficient starting material strategies has continued this past year in two different areas. In the first, we have explored the prospect of using a parallel computer in the graph matching process described in the following section and in the second we have developed a solution to the class IV problem (see above) which is described in a subsequent section.

C.1.c Subgraph Search Using Parallelism.

Subgraph matching is an important method used in many different computer applications in organic chemistry, including the recognition of functional groups, synthesis planning, constraint testing in structure generation, selection of starting materials for synthesis, and structure oriented retrieval. The fundamental problem is, given a ***query substructure*** (QS) and a ***candidate superstructure*** (CS), determine if there exists a mapping of the atoms (nodes) of the substructure onto the candidate superstructure such that the connected atom pairs in the query substructure are also connected in the superstructure, and that the atom and bond types also correspond.

Although substructure search is a non-numerical problem, it is computationally demanding because ultimately it involves establishing an atom by atom correspondence between the QS and the CS, and this problem is a member of the class of NP-complete problems. In a worst case for **N** atoms in the QS and **M** atoms in the CS (**M>N**), one may have to consider **N!/(M-N)!** mappings for each CS. The objective of our research was to explore the feasibility of applying parallel processing to this problem.

Although the node matching process is an NP-complete problem, if we eliminate all backtracking, the order of the algorithm reduces to **O(N)**, where **N** is the number of

atoms in the subgraph. This would represent a major improvement. Unfortunately, the algorithm is now NP-complete with respect to sequential processors.

We proposed a "star" configuration architecture; a central processor with a communication line to a number of lower processors, with no direct communication allowed between the lower processors. Each processor has a small amount of memory as a "working space", which avoids the problems inherent in shared memory. The communication packets are compact to reduce the storage and communication burden on the central processor.

The simulation algorithm called MOLSIM was implemented using the SIMULA language. We studied this algorithm as a function of the number of processors and the nature of the particular graphs being matched (6-31 non-H atoms). We found an average utilization of 84% on a 25 cpu machine (figured as total processor time/real time), but only 60% on a 50 cpu machine although for some structure matching questions, the efficiency reached 97%. The average speed enhancement using this size machine (50 processors) was a factor of 30. A real machine of the architecture needed to run this algorithm exists at Purdue University, and time is being requested to test the algorithm in real time. This algorithm is a unique approach to the problem of graph matching and will likely become practical when parallel processors are commonly available and inexpensive. (This work is submitted for publication in J. Chem. Inf. Comput. Sci..) *C.1.b MCS-- Maximal Common Subgraph Search.*

The second area of starting material strategy work this year has been in solving the class IV problem given above. Our solution to this problem involves development of a new efficient maximal common subgraph matching algorithm. Since chemists represent organic molecules as graphs, computational chemists need graph theoretical techniques such as graph isomorphism, subgraph search, and maximal common subgraph search. Of these three important procedures, maximal common subgraph search (MCSS) remains the most difficult and least utilized.

The extensive computational demands of MCSS has restricted its possible uses. We have previously noted that maximal common subgraph search could be useful in our starting material selection program, SST, but that the computational demands were too rigorous.* Cone et al. who has used MCSS in their "self-training interpretive and retrieval system" (STIRS), has noted that other potential uses of for MCSS include computer-assisted organic synthesis and structure activity studies.**

Given two graphs, finding a ***common subgraph*** involves discovering the assignment of some of the nodes and edges of one graph onto the other graph while preserving the adjacency relationships of the nodes. The ***size*** of the common subgraph is the number of edges preserved in the assignment. If there exists no common subgraph of larger size, the common subgraph is called ***maximal.***

Our approach to MCSS was to reduce redundant searching and try to shrink the size of the search space. We observe that most libraries of chemicals have compounds which have similarities; by capitalizing on these similarities we might be able to reduce the search space. The essential principle is that if we know a relationship between library graph A and B we can relate query graph C to A, then we may therefore already know something about the relationship of C to B. We establish the relationship between A and B in a one-time-only preprocessing of the library.

---

*W.T. Wipke, D. Rogers, *J. Chem. Inf. Comput. Sci.*, 1984, ,(in press)

**M.M Cone, R. Venkataraghavan, F.W. Mclafferty, *J. Am. Chem. Soc.*, *99*, 7668, (1977)

The objective of our study was to demonstrate the feasibility and study properties of such an algorithm. The following design features were deemed important.

- The processed library should be storage efficient.
- It should be possible to add compounds ***incrementally*** to the file, to avoid the cost of reprocessing the file whenever changed.
- The time required to store a new compound in the processed library should be minimal; preferably, an upper limit on this time requirement should be known.
- It should be possible to create the compound from its processed form, so that both processed and non-processed libraries do not have to be kept.

We wished the system to be useful for a range of non-preselected queries, therefore we ruled out simply "training" the system for a small class of query choices or depending on detailed knowledge of the allowed queries.

We had noticed that identical common subgraph candidates are often generated during the search of different compounds against a query. We envisioned collecting these common segments into an intermediate graph, to allow the search to be conducted once over the common feature. If our search procedure can take advantage of the abstraction of the common graph from two or more test compounds, then the number of attempted subgraph matches will be reduced.

A non-recursive FORTRAN algorithm was implemented for the MCS-1 program using a tree-structured storage file. The tree storage search algorithm gives reductions from 70 to 90% in the search relative to conventional unstructured sequential storage systems. Reductions are especially good for the case where the library contains smaller graphs or a series of similar graphs. The general structure of the search tree is determined early in its creation; once the major nodes of the tree are established, addition of compounds rarely alters it significantly. Sorting experiments showed that the tree can be "seeded" in such a way that improved results can be obtained when searching over the seeded library relative to the unseeded library.

We have applied this MCS-1 algorithm to the Class IV starting material recognition problem. The initial abstraction of the starting material library resulted in an abstracted library of significantly reduced size. Organizing this abstracted library in a tree-structured form allows the discovery of starting materials which do not have a subgraph or super graph relationship with the target. A trial run with morphine was successful in pointing out an interesting starting material candidate not found by the SST program. This work is being submitted to the J. Chem. Inf. Comput. Sci..

C.2 XENO Program Developments

The metabolic fate of various compounds in the human body is extremely complex, yet extremely important for it is known that through metabolism certain otherwise harmless compounds are converted into toxic and possibly carcinogenic agents. Because of this complexity it is difficult, looking at a given compound, to forecast potential biological activity of that given compound. The objective of this proposal is to develop a practical computer program by which a biochemist or metabolism expert can explore the metabolites of a given compound and be alerted to the plausible biological activity of each metabolite.

This research aims to explore the degree to which current knowledge of metabolism can be used by a computer program to make reasonable projections of what metabolites

might result from exposure of a compound to a biological system. The project involves representing in a computer the metabolic processes with all known specificities and applying these processes to a xenobiotic compound to generate a set of plausible metabolites which may themselves be further metabolized. We also plan an evaluation module to appraise the plausible biological activities of each metabolite using a rule base to relate chemical structure to biological activity. Thus the attention of the experimentalist will be attracted to those metabolites that are likely to be biologically active.

C.2.a Atomic Charges.

During the past year, we implemented rapid atomic charge calculation algorithms to be used with the XENO program in order to better predict the biological activity of metabolites. The algorithms were based on Gasteiger's PEOE (partial equalization of orbital electronegativity)* and SD-POE (sigma dependent pi orbital electronegativity)** models. The PEOE model has been used for sigma charge calculations for sigma bonded and non-conjugated pi systems. The SD-POE model has been used for conjugated aliphatic and single ring aromatic molecules. For polyaromatic systems, the pi charge calculations are being implemented using the SD-POE model and the Longuet-Higgins approximation.*** In the PEOE model, the SD-POE model, and polyaromatic hydrocarbon pi charge calculation, atoms are characterized by their orbital electronegativities.

We have shown that the charges so calculated are reasonable when compared to the work of others in the literature. Our purpose is then to correlate the biological activity of metabolites with the atomic charge on electrophilic centers in the metabolites. We also think that metabolism itself can be controlled to some extent by atomic charges so the metabolic transforms may make use of this data eventually.

C.2.b pKa Calculations.

We have continued work on the problem of the estimation of the dissociation constants for organic acid and bases to be used to increase the expertise of the XENO program. We have been investigating two approaches to do this type of estimation: LFER (linear free energy relations), and theoretical or quantum chemical approaches.****

The LFER computation is performed automatically by first selecting appropriate skeleton structures from a library, then recognizing attached groups and finally calculating the pKa from the relevant equations and group substituent constants. This is the first automatic pKa estimation algorithm ever developed and promises wide utility on its own outside of the XENO program.

To determine the most representative pKa if more than one acid or base center is present, several empirical rules are followed:

- **acids** - use the ionized form of the acid center with the lowest pKa as a substituent for the pKa calculation of the next lowest acid, and so on.

---

* J. Gasteiger and M. Marsili, *Tetrahedron lett.*, *34*, 3181, (1981)

** J. Gasteiger and M. Marsili, *Tetrahedron*, *36*, 3219, (1980)

*** H, C. Longuet-Higgins, *J. Chem. Phys.*, *18*, 275, (1950)

**** D.D. Perrin, B. Dempsey, and E.P. Serjeant, "pKa Prediction for Organic Acids and Bases", Chapman and Hall, New York, 1981.

- **bases -** use the protonated form of the base center with the highest pKa as a substituent for the pKa calculation of the next highest base, and so on.
- **strong acids and bases -** compute pKa's for the acids first, then use ionized form of the acids as substituents to compute pKa's of the bases.
- **strong acids and weak basis -** compute pKa's for bases first, then use protonated form of the bases as substituents to compute pKa's of the acids.
- **weak acids and weak basis -** compute pKa's for weak bases first, then use deprotonated form of bases as substituents to compute pKa's for the acids.

C.2.g Collaborative Efforts. In the past year most work was aimed at program development rather than application to laboratory problems, however in the next year we do expect after completion of the current program modifications to perform several practical analyses in conjunction with our collaborators at ICI of UK, NIH, and other parties that have indicated interest.

The SECS project continues to have collaborations with the pharmaceutical industry who are adding chemical transforms and doing some joint program development, for example, Dr. Yanaka continued work started at Santa Cruz after he returned to Kureha Chemical in Japan and a paper has been prepared on that work.

D. List of Current Project Publications

1. Wipke, W.T., and Rogers, D.: ***Rapid Subgraph Search Using Parallelism*** J. Chen. Inf. Comput. Sci (submitted 24 April 84)
2. Wipke, W.T.: "An Integrated System for Drug Design" in *Computers A-Z: A Manufacturer's Guide to Hardward and Software for the Pharmaceutical Industry* Aster Publishing Co., Sprinfield, Oregon, (in press)
3. Wipke, W.T. and Huber, M.: ***Symmetry and organic synthetic design.*** Accepted in Tetrahedron.
4. Wipke, W.T., Ouchi, G.I. and Chou, J.T.: *Computer-Assisted Prediction of Metabolism.* IN L. Goldberg (Ed.), STRUCTURE-ACTIVITY CORRELATIONS AS A PREDICTIVE TOOL IN TOXICOLOGY. Hemisphere Publishing Corp., New York, 1983.
5. Johnson, C.K., Thiessen, W.E., Burnett, M.N., Condran, P. Ronlan, A., Yanaka, M. and Wipke, W.T.: ***Systematic derivation of chemical procedures for transforming surplus hazardous chemicals to useful products, J. of Hazardous Materials.*** (In press)
6. Dolata, D.P.: *QED: Automated Inference in Planning Organic Synthesis* (Ph.D. dissertation). University of California, Santa Cruz, 1984.
7. Rogers, D.: *Artificial Intelligence in Organic Chemistry. SST: Starting Material Selection Strategies* (Ph. D. dissertation). University of California, Santa Cruz, 1984.

Wipke, W.T., and Rogers, D.: *Artificial Intelligence in Organic Synthesis. SST: Starting Material Selection Strategies. An Application of Superstructure Search.* J. Chem. Inf. Comput. Sci., 24:0000, 1984.

E. Funding Status

1. Computer-Assisted Prediction of Xenobiotic Metabolism
   Principal Investigator: W. Todd Wipke, Professor, UCSC
   Agency: NIH, Environmental Health Sciences; No: ES02845-02
   4/1/82-3/31/85 $257,801 TDC
   4/1/84-3/31/85 $ 89,140 TDC

2. Graphical Display of Chemical Inferences and Molecular Relationships
   Principal Investigator: W. Todd Wipke, Professor, UCSC
   Agency: Evans and Sutherland Corporation
   Gift of PS300 B/W High Performance Graphical Display System
   Permanent, value $95,000 TDC

3. Computer Synthesis
   Principal Investigator: W. Todd Wipke, Professor, UCSC
   Agency: Stauffer Chemical Company
   Permanent, $6,000 TDC

F. Research Environment

At the University of California, Santa Cruz, we are connected to the SUMEX-AIM resource by a 4800 baud multiplexed leased line. Our video terminals consist of a Z-29, DM-3025, TI745, CDI-1030, DIABLO 1620, and an ADM-3A. We have a PS300 graphic display which is driven by SUMEX. UCSC has only a small IBM 370/145, a PDP-11/45, 11/70 and a VAX 11/780, (the 11's are restricted to running small jobs for student time-sharing) all of which are unsuitable for our current research. The SECS laboratory is located in 125 Thimann Laboratories, adjacent to the synthetic organic laboratories at Santa Cruz.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination 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 graphic terminal equipment at Santa Cruz, so much more convenient for structure input and output. Demonstrations and sample synthetic analyses were generated for numerous visitors from the US and abroad. Demonstrations of SECS in Sweden were performed by Dr. R. E. Carter, University of Lund, Sweden, at many universities and companies.

Professor Wipke has also used several SUMEX programs such as CONGEN in his course on Computers and Information Processing in Chemistry. Communication between SECS collaborators is facilitated by using SUMEX message drops, especially when time differences between the U.S. and Europe and Australia makes normal telephone communication difficult. Testing and collaboration on the XENO and FSECS project with researchers at the NCI depend on having access through SUMEX and TYMNET.

Collaboration with Lund University. The introduction of SECS to organic chemists in Sweden was one of the seeds that led to the establishment of a computer graphics laboratory for organic chemistry at the University of Lund, with strong support from a government agency, the National Swedish Board for Technical Development (STU).

Interest in the applications of computers and computer graphics in organic chemistry has spread very rapidly throughout the country, and chemists at all of the major Swedish universities, as well as in the pharmaceutical industry, have taken steps to participate in the exciting developments in this field.

Interest in the pedagogical value of SECS at the graduate level has led to its use to illustrate the concepts of retrosynthetic planning and analysis in conjunction with a course given by Prof. Paul Helquist (SUNY, Stony Brook): "Synthetic Organic Chemistry---Modern Methods and Strategy". The same course was given at the Royal Institute of Technology in Stockholm, and at the Chemical Center in Lund, using SECS as an integral part of the course.

A Workshop on Computers in Organic Chemistry was sponsored by STU on May 17-18,1983, in Gothenburg to help organic chemists in Sweden enter this area of work. Daniel P. Dolata from Prof Wipke's group in SC was an invited speaker.

A chemist from Lund, Dr. Alvin Ronlan, spent a sabbatical leave with the Wipke group, and a graduate of the Wipke group, Dr. Dolata, is spending a postdoctoral stay in Lund.

In collaboration with the SECS group at UCSC, Dolata will install the SECS program and the QED program on a new VAX 11/780 for the use of the chemists in Lund, and will continue research with QED. For example, it would be of interest to develop rule bases to assist the chemist in structure elucidation, and structure-activity relationships.

Another area of collaboration involves compilation of chemical transforms by the chemists in Lund. Some of the chemists in Lund work with natural products (isolation and synthesis), with a view toward the discovery and characterization of physiologically active substances. For example a strongly mutagenic compound has been isolated from a Swedish mushroom (Lactarius vellereus), its structure determined, and a total synthesis elaborated. In other work, a traditional abortifacient from Bangladesh is being isolated from plant material, and a psychoactive substance is being extracted from the leaves of a Nigerian plant. A collaboration with a university in Holland is now developing along similar lines, and Cornell University is planning a similar center for computer applications in chemistry.

B. Examples of Cross-fertilization with other SUMEX-AIM Projects

The AILIST bulletin board has been used extensively for interacting with many projects and locating references for further information related to program design and AI technology. There are no longer any other chemical or biochemical projects on SUMEX so our interaction with the community is limited to AI technology interchange, attending seminars at Stanford, etc.

C. Critique of Resource Services

SUMEX-AIM gives us at UCSC, a small university, the advantages of a larger group of colleagues, and interaction with scientists all over the country. Previously we were provided very good service by SUMEX-AIM, but since 1 April 1984, the computer service has been very poor. Although the National AIM usage of SUMEX has been small, our project has been put in a separate class with a 3% cpu limitation. This is a very severe restriction which prevents short usage peaks from being averaged with other users. Our project is the only project subjected to such limitations. The poor response time we are observing (load averages of 25-50!) is significantly hindering our ability to perform the research NIH funded to be done on SUMEX. This is worsened by the fact we are in the last year of the project.

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

SECS 2.9 has been installed on the CompuServe computer networks for the past three years so anyone can access it without having to convert code for their machine. This has proved very useful as a method of getting people to experiment with this new technology. Dr. George Purvis of Battelle has accessed SECS via CompuServe, as has Gene Dougherty of Rohm and Haas and many others. SECS also resides on the Medicindat machine at the University of Gothenborg, Sweden, and is available all over Sweden by phone. Similarly in Australia, SECS resides at the University of Western Australia and is available throughout Australia over CSIRONET. A lecture series was given on SECS in Tokyo and SECS has been installed at two locations in Japan. FSECS has been installed on a DEC-10 at Oak Ridge National Laboratory and serves for collaborative development of that approach with Carroll Johnson. PRXBLD has been disseminated to over 30 sites on various types of computers including DEC-10, DEC-20, IBM, VAX, PRIME and Honeywell.

III. RESEARCH PLANS (4/84-4/85)

A. Near-Term Project Goals and Plans Our research projects will move off SUMEX-AIM by 31 March 1985 to some other as yet unspecified computer system. Therefore our research objectives on SUMEX-AIM are to complete research in progress, consolidate programs and files for moving to another system. The QED and SST projects have been completed and the first phase of the RXAN project outlined last year has been completed. On the SECS project, the reaction library is being extended by Dr. Iwataki. We will continue to collaborate with coworkers in SECS research on other machines but on SUMEX will primarily be preparing SECS for removal from SUMEX. We are exploring PROLOG as a replacement for the QED system and plan some preliminary sample PROLOG programs to compare the capabilities of PROLOG and QED. But the majority of our activities will be aimed at completion of the XENO program.

XENO Goals Our objectives for this year follow our plan in the original proposal. Basically in this next year we plan to complete the implementation of algorithms not yet completed and focus on testing with applications to demonstrate the current power of XENO on typical laboratory metabolism problems. In this last year of this project our goal is to bring XENO to a relatively stable finished point which will be useful to other researchers. We believe we will complete the algorithms in progress, document them and submit publications on all of the work within the year. The major areas of focus are listed below.

A.1 Atomic Charge Calculations. We also plan to complete our correlations between atomic charges and sites of metabolism, as we have already done with bond reactivity. When such correlations are established, then they can be used to guide XENO to apply metabolic transforms more selectively to the most active parts of the molecules.

A.2 pKa Calculations. We plan to complete testing of the pKa algorithm on different groups on metabolites so that information may be used by XENO for activity evaluation, selection of further possible metabolism, and estimation of excretion and transport.

A.3 Three-Dimensional Criteria. We plan to complete our work on three-dimensional constraints that apply to metabolism which have been obtained through study of many metabolic studies in the literature. This will require extensions of the

ALCHEM language to accommodate new types of three-dimensional relationships such as overall molecular size, width, thickness, ratio of length to width, etc.

A.4 Log P Calculation. In order to more accurately estimate the possibility of excretion and binding for metabolism, we plan to incorporate a log P calculation module. This will provide the partition coefficient between octanol and water. There are already programs to calculate log P, and these have been shown to be very accurate. Our objective, time permitting, will be to include such a module in the XENO program and correlate log P with metabolism transform application and with excretion.

B. Justification and Requirements for Continued Use of SUMEX

The XENO project which resides on SUMEX is in its last year of support, consequently we need to complete that research on the SUMEX machine. By 31 March 1985, we plan to move XENO and all our research off SUMEX onto some other computer. We are currently exploring what machine may be suitable and available. After 1 April 1985 we will not need SUMEX for computational support, but will need access to be able to retrieve certain files from archive, respond to electronic mail, and continue to participate in the AIM scientific interchange through electronic mail and bulletin boards. It is not practical to retrieve every file we have ever archived, it would use too much SUMEX operator time, and it is unnecessary as long as we can access them if we need them in the future. That access would not require significant resources.

However prior to 31 March 1985 we have obligations to complete the research on the XENO project supported by NIH and need sufficient SUMEX cpu time to accomplish this goal. This means normal editing, compile, load, and test executions plus some application runs to some metabolic problems. It appears the current removal of our project from the National AIM portion of the SUMEX-AIM resource and placement in a class restricted to 3% peak utilization is hindering the research productivity of this project. We are experiencing load averages of 25-50 a high percentage of the time. We request to have our project placed back in the National AIM portion of the SUMEX-AIM resource as we were allocated, and we will carefully monitor to see that our resource utilization does not exceed our quota of time. We feel this is a reasonable request in light of the mission of SUMEX-AIM to the National community of which this project is a part.

C. Needs Beyond SUMEX-AIM As mentioned above our project needs additional computing resources and we are exploring acquiring a computer for installation at UCSC and obtaining the necessary resources to support it. We are seeking information about comparisons between machines and cost effectiveness of different hardware combinations.

D. Recommendations for Community and Resource Development

It appears the SUMEX-AIM resource is increasingly becoming basically a Stanford resource and that there is a difference between the portion of the resource allocated to the National community and the portion actually used by the National community. Our project is part of the National community and in need of better service, we hope that can be improved.

An important part of medicine is treatment of diseases with drugs, chemicals, chemicals that were designed and synthesized by chemists. Since the termination of the DENDRAL project, there seems to be declining support for artificial intelligence applications in chemistry. We feel that support of this area is essential to the advancement of medicine in this country. The lack of chemists on NIH Research Resources computing peer review is contributing to the problem. In general the AIM community would benefit by involving disciplines other than computer science.

II.A.2.5. SOLVER Project

SOLVER: Problem Solving Expertise

Dr. P. E. Johnson
Center for Research in Human Learning
University of Minnesota

Dr. W. B. Thompson
Department of Computer Science
University of Minnesota

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

This project focuses upon the development of strategies for discovering and documenting the knowledge and skill of expert problem solvers. In the last fifteen years, considerable progress has been made in synthesizing the expertise required for solving extremely complex problems. Computer programs exist with competency comparable to human experts in diverse areas ranging from the analysis of mass spectrograms and nuclear magnetic resonance (Dendral) to the diagnosis of certain infectious diseases (Mycin).

Design of an expert system for a particular task domain usually involves the interaction of two distinct groups of individuals, "knowledge engineers," who are primarily concerned with the specification and implementation of formal problem solving techniques, and "experts" (in the relevant problem area) who provide factual and heuristic information of use for the problem solving task under consideration. Typically the knowledge engineer consults with one or more experts and decides on a particular representational structure and inference strategy. Next, "units" of factual information are specified. That is, properties of the problem domain are decomposed into a set of manageable elements suitable for processing by the inference operations. Once this organization has been established, major efforts are required to refine representations and acquire factual knowledge organized in an appropriate form. Substantial research problems exist in developing more effective representations, improving the inference process, and in finding better means of acquiring information from either experts or the problem area itself.

Programs currently exist for empirical investigation of some of these questions for a particular problem domain (e.g. AGE, UNITS, RLL). These tools allow the investigation of alternate organizations, inference strategies, and rule bases in an efficient manner. What is still lacking, however, is a theoretical framework capable of reducing dependence on the expert's intuition or on near exhaustive testing of possible organizations. Despite their successes, there seems to be a consensus that expert systems could be better than they are. Most expert systems embody only the limited amount of expertise that individuals are able to report in a particular, constrained language (e.g. production rules). If current systems are approximately as good as human experts, given that they represent only a portion of what individual human experts know, then improvement in the "knowledge capturing" process should lead to systems with considerably better performance.

B. Medical Relevance and Collaboration

Collaboration with Dr. James Moller MD in the Department of Pediatrics, Dr. Donald Connelly MD in the Department of Laboratory Medicine, at the University of Minnesota. Collaboration with Dr. Eugene Rich MD and Dr. Terry Crowson MD at St. Paul Ramsey Medical Center.

C. Highlights of Research Progress

Accomplishments of This Past Year -- Prior research at Minnesota on expertise in diagnosis of congenital heart disease has resulted in a theory of diagnosis and an embodiment of that theory in the form of a computer simulation model, ***Galen***, which diagnoses cases of congenital heart disease [Thompson, Johnson & Moen, 1983].

Galen is descended from two earlier programs written here at Minnesota: *Diagnoser* and *Deducer* [Swanson, 1977]. Deducer is a program that builds hemodynamic models of the circulatory system that describe specific diseases. The models are built by using knowledge about how idealized parts of the circulatory system are causally related. Diagnoser is a recognition-driven program that performs diagnoses by successively hypothesizing one or more of these models and matching them against patient data. The models that match best are used as the final diagnosis. A series of experiments carried out at Minnesota have shown that Diagnoser/Deducer performs as well (and sometimes better) than expert human cardiologists [Johnson et al., 1981].

Despite their early successes, Diagnoser and Deducer did not have a clear, comprehensible structure that is required for the kind of experiments we wish to perform. Galen was built to remedy this problem, taking advantage of the experience gained in the design of Diagnoser and Deducer.

Galen consists of four major components: a working memory called the ***scratchpad***, a knowledge base of rules and hypotheses, a procedure called the ***proposer*** and a procedure called the ***reviewer***.

The scratchpad contains data about the problem that Galen is trying to solve and the hypotheses that are being investigated to explain that data. In effect, the scratchpad represents Galen's current execution state.

Rules are pattern-action pairs. The pattern part of a rule describes a possible state of the scratchpad. Patterns can contain imbedded logical connectives (e.g. ANDs, ORs, NOTs) and can be constructed to match at varying levels of detail. The action part is a procedure that is executed if the pattern part matches the scratchpad's contents. Each action part writes an assertion on the scratchpad about a hypothesis, together with the evidence for making that assertion. These assertions can express that a new hypothesis is being considered, or that an old hypothesis has been accepted, rejected, confirmed or disconfirmed. Action parts can also assert that a hypothesis is sufficient to solve the current problem, or that the problem is not solvable.

Because the pattern parts of rules can examine anything on the scratchpad, it is possible to express rules about hypotheses as well as rules about problem data. In particular, this makes it possible to directly examine the accumulated evidence for and against each currently contending hypothesis, making numerical measures of certainty unnecessary.

A hypothesis is simply a named collection of rules. The hypotheses in Galen's knowledge base can be thought of as a directed graph, in which vertices are hypotheses and edges are rules. One hypothesis "points to" another if the first hypothesis contains a rule whose action part can assert something about the second.

The level of detail of such a knowledge base leads to serious problems with the computational complexity of search processes. Galen focuses its computational resources so that the knowledge embodied in the graph of hypotheses can be used in an efficient manner. Successful diagnoses result from good first hypotheses about possible defects and efficient mechanisms for refining these hypotheses.

Galen works by using the proposer and the reviewer to investigate hypotheses (i.e. search the graph) by applying rules (i.e. following the edges from one vertex to another). Whenever a new piece of problem data is written on the scratchpad, the proposer applies all relevant rules specific to the type of that piece of data. These rules write assertions on the scratchpad about new hypotheses, effectively identifying vertices in the graph that are worthy starting points for further search. Next, the reviewer applies all relevant rules contained in the hypotheses that are named in assertions on the scratchpad. Successfully applying one of these rules corresponds to propagating the search along a specific edge of the graph. The search is constrained because (1) only the most promising vertices in the graph are ever used to initiate search; (2) only a small number of edges are ever followed; and (3) most rules in a hypothesis deal with evidence for and against the hypothesis itself, giving a graph where the number of effectively outward-pointing edges at each vertex is small.

The read-propose-review cycle repeats in this way until some hypothesis has been shown correct, until the problem has been shown unsolvable, or until all the data has been examined.

Currently, data given to Galen is taken from a (possibly imaginary) patient's medical chart. Hypotheses in the knowledge base represent the ten most commonly occurring congenital heart diseases and their variants, useful intermediate physiological findings, and classes of hypotheses. Since hypotheses are implemented as named teams of production rules, it is also possible to represent other kinds of hypotheses should the need arise. Moreover, Galen has been constructed so that its inference engine does not contain any procedures specialized for pediatric cardiology. It is therefore conceivable to extend Galen to other domains if effective knowledge bases for those domains can be constructed.

To determine the generality of our model of expertise in diagnostic reasoning, we are also investigating domains outside medicine. As part of this effort, we have developed a computational model of the fault localization process in program debugging [Sedlmeyer, 1983] that is not based directly on Galen. As with our work in congenital heart disease, we have concentrated on the design of mechanisms for structuring problem specific knowledge and for focusing limited computational resources.

Research in Progress -- Since human experts are notoriously poor at describing their own knowledge, our work requires the creation of problem solving tasks through which experts can reveal criteria for initiating specific hypotheses and methods for investigating those hypotheses.

Current techniques of representing hypotheses and their expectations for diagnosis do not, however, provide much detailed information about the control processes experts use to guide their reasoning. Such control processes typically incorporate highly refined heuristics about which the experts are almost wholly unaware. To discover the needed control knowledge, we ask experts to complete tasks in which we have systematically perturbed aspects of the problem data. The data in these tasks are chosen so that members of an overlapping set of hypotheses will be suggested during while solving the problem. Success in solving such problems depends on the ability to overcome an initially plausible incorrect hypothesis in favor of a later, more correct alternative.

Several examples illustrate our approach. We are studying performance of Galen on "garden path" cases [Johnson & Thompson, 1981] that were initially misdiagnosed in hospital files. Analysis of such cases suggest that errors are made because experts rely on very efficient heuristics that are not universally correct. In one such example, a seemingly plausible hypothesis is suggested early in the case. Although the hypothesis superficially seems to explain what is observed about the patient, the hypothesis is incorrect. Because the incorrect hypothesis seems marginally adequate, it acts to prevent a more correct hypothesis from being suggested in its place. Success in such a case hinges on the ability to use the proper set of competing hypotheses in order to provide more than one explanation of the case data. Investigation of this phenomenon in human experts has suggested implementation of "transition rules" linking disease hypotheses in Galen. It has also suggested implementation of "monitor hypotheses" that watch for potential garden path errors and avoid them before they become serious.

We are also investigating several research questions relevant to the architecture of Galen. We have designed an interface to Galen so that users who are unfamiliar with the inner workings of the program can interactively enter case data. Designing the interface raised questions about what forms of data are necessary to adequately and completely represent all possible cases. We are also studying ways in which a causal reasoning component can be integrated with the prototypical reasoning components (the Proposer and Reviewer) that are already present in Galen. In particular, we are interested in studying ways in which causal reasoning can aid or replace prototypic reasoning when it becomes inadequate to reach a diagnosis.

In another project, we are investigating methods of probabilistic reasoning. Most systems rely on numerical schemes for weighting evidence or ranking observed data. These weights are often probabilistic in nature, but other schemes have also been used. Mycin, for example, uses certainty factors and PIP uses likelihoods composed of matching scores and binding scores. In contrast, humans do not seem to rely upon such numerical techniques. Research has shown that people are often quite poor at probabilistic reasoning. However, experts make decisions which involve weighting evidence and selecting from competing alternatives. They must utilize a reasoning process which serves as an alternative to a numerical weighting technique.

We believe the process of weighting alternatives along various criterial dimensions is not a domain specific technique, but rather a general process which is applied in specific instances. In the coming year, we will examine this process in various domains and attempt to utilize the results in designing more powerful reasoning techniques.

In the area of law, our work has focused on the area of corporate law (the problem of structuring a proposed corporate acquisition). We have collected data from 24 practicing lawyers and in the coming year a PhD thesis will be completed describing this work. In the coming year we will be also completing a study of the corporate acquisition problem in management in order to further refine our knowledge capturing tools.

D. List of Relevant Publications

1. Connelly, D. and Johnson, P.E.: ***Medical problem solving.*** Human Pathology, 11(5):412-419, 1980.

2. Elstein, A., Gorry, A., Johnson, P. and Kassirer, J.: ***Proposed Research Efforts.*** IN D.C. Connelly, E. Benson and D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 327-334.

3. Feltovich, P.J.: *Knowledge based components of expertise in medical diagnosis.* Learning Research and Development Center Technical Report PDS-2, University of Pittsburgh, September, 1981.

4. Feltovich, P.J., Johnson, P.E., Moller, J.H. and Swanson, D.B.: *The Role and Development of Medical Knowledge in Diagnostic Expertise.* IN W. Clancey and E.H. Shortliffe (Eds.), READINGS IN MEDICAL AI. (In press)

5. Johnson, P.E.: *Cognitive Models of Medical Problem Solvers.* IN D.C. Connelly, E. Benson, D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 39-51.

6. Johnson, P.E.: *What kind of expert should a system be?* J. Medicine and Philosophy, 8:77-97, 1983.

7. Johnson, P.E., *The Expert Mind: A new Challenge for the Information Scientist* IN Th. M. A. Bemelmans (Ed.), INFORMATION SYSTEM DEVELOPMENT FOR ORGANIZATIONAL EFFECTIVENESS, Elsevier Science Publishers B. V. (North-Holland), 1984.

8. Johnson, P.E., Severance, D.G. and Feltovich, P.J.: *Design of decision support systems in medicine: Rationale and principles from the analysis of physician expertise.* Proc. Twelfth Hawaii International Conference on System Science, Western Periodicals Co. 3:105-118, 1979.

9. Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic reasoning.* Cognitive Science 5:235-283, 1981.

10. Johnson, P.E. and Hassebrock, F.: *Validating Computer Simulation Models of Expert Reasoning.* IN R. Trappl (Ed.), CYBERNETICS AND SYSTEMS RESEARCH. North-Holland Publishing Co., 1982.

11. Johnson, P.E. and Thompson, W.B.: *Strolling down the garden path: Detection and recovery from error in expert problem solving.* Proc. Seventh IJCAI, Vancouver, B.C., August, 1981, pp. 214-217.

12. Johnson, P.E., Hassebrock, F. and Moller, J.H.: *Multimethod study of clinical judgement.* Organizational Behavior and Human Performance 30:201-230, 1982.

13. Moller, J.H., Bass, G.M., Jr. and Johnson, P.E.: *New techniques in the construction of patient management problems.* Medical Education 15:150-153, 1981.

14. Swanson, D.B.: *Computer simulation of expert problem solving in medical diagnosis.* Unpublished Ph.D. dissertation, University of Minnesota, 1978.

15. Swanson, D.B., Feltovich, P.J. and Johnson, P.E.: *Psychological Analysis of Physician Expertise: Implications for The Design of Decision Support Systems.* In D.B. Shires and H. Wold (Eds.), MEDINFO77, North-Holland Publishing Co., Amsterdam, 1977, pp. 161-164.

16. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

17. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: ***Knowledge-based fault localization in debugging***. The Journal of Systems and Software (in press).

18. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: ***Diagnostic reasoning in software fault localization***. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

19. Sedlmeyer, R.L., Thompson, W.B., & Johnson, P.E.: ***Knowledge-Based Fault Localization in Debugging***. The Journal of Systems and Software (in press).

E. Funding and Support

Work on the SOLVER project is currently supported by a grant from the Control Data Corporation to Paul Johnson ($90,000; 1983-85) and by a grant from the Microelectronics and Information Sciences Center at the University of Minnesota to Paul Johnson, William Thompson and two colleagues ($800,000; 1984-87).

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

Work in medical diagnosis is carried out with the cooperation of faculty and students in the University of Minnesota Medical School and St. Paul Ramsey Medical Center.

B. Sharing and Interactions with Other SUMEX-AIM Projects

A year ago, conversations were begun with William Clancey at Stanford University regarding collaboration on the study of current knowledge capturing methods. We plan to develop this collaboration in the coming year.

C. Critique of Resource Management

(None)

III. RESEARCH PLANS

A. Project Goals and Plans

Near term -- Our research objectives in the near term can be divided in three parts. First, we are committed to the design, implementation, and evaluation of Galen, as described above. We have completed an interactive front end so that physicians can directly enter patient data, and Galen's knowledge base is currently being "tuned" with the help of Dr. James Moller MD, an expert physician collaborator from the University of Minnesota Pediatric Cardiology Clinic. During the coming year, Galen's performance will be compared with that of the Diagnoser program and with expert physicians.

Our second objective consists of making extensions to the knowledge capturing strategies developed in our original work in medical diagnosis. In the near term this work will examine descriptive strategies in which experts attempt to use a formalized language to express what they know (e.g. production rules), observational strategies in which experts perform tasks designed to reveal information from which a theory of task specific expertise can be built, and intuitive strategies in which either experts behave as knowledge engineers or knowledge engineers attempt to perform as pseudo experts. In the coming year we will also be attempting to develop a program to automate the early stages of

knowledge capturing, analogous to the "prototype stage" of design referred to in software engineering.

Our third near term objective will be to investigate one of the central problems of recognition based problem solving, how to classify problems when solving them. Questions related to problem classification which we will be examining include: What patterns do experts and novices detect in a problem that allows them to classify it as an instance of a problem type that is already known? How does an expert make an initial choice of the level of abstraction to be used in solving a problem? How can an expert recover from an initial incorrect choice of levels? How can the difference between causal and prototypic modes of reasoning be modeled as differences in levels of abstraction, and how can a common model for these two types of reasoning be constructed? We will be pursuing these questions in the area of physics problem solving, as well as in medicine.

Long range -- Our long range objective is to improve the methodology of the "knowledge capturing" process that occurs in the early stages of the development of expert systems when problem decomposition and solution strategies are being specified. Several related questions of interest include: What are the performance consequences of different approaches, how can these consequences be evaluated, and what tools can assist in making the best choice? How can organizations be determined which not only perform well, but are structured so as to facilitate knowledge acquisition from human experts? In the coming year we will be exploring these questions in areas of design and management as well as in law, physics and medicine.

B. Justification and Requirements for Continued SUMEX Use

Our current model development takes advantage of the sophisticated Lisp programming environment on SUMEX. Although much current work with Galen is done using a version running on a local VAX 11/780, we continue to benefit from the interaction with other researchers facilitated by the SUMEX system. We expect to use SUMEX to allow other groups access to the Galen program. We also plan to continue use of the knowledge engineering tools available on SUMEX.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

Paul Johnson is a member of the group of investigators who have recently submitted a proposal to establish a national computer network for cognitive scientists (COGNET). In addition, our current grant will permit the purchase of some single-user computers (we are currently comparing several alternative machines). SUMEX will continue to be used for collaborative activities and for program development requiring tools not available locally.

D. Recommendations for Future Community and Resource Development

(None)

II.A.3. Pilot Stanford Projects

Following are descriptions of the informal pilot projects currently using the Stanford portion of the SUMEX-AIM resource, pending funding, full review, and authorization.

In addition to the progress reports presented here, abstracts for each project are submitted on a separate Scientific Subproject Form.

II.A.3.1. CAMDA Project

CAMDA Project

CAMDA Research Staff:

Samuel Holtzman, Co-PI	**Engineering-Economic Systems**
Prof. Ronald A. Howard, Co-PI	**Engineering-Economic Systems**
Jack Breese	**Engineering-Economic Systems**
Dr. Emmet Lamb	**School of Medicine**
Dr. Robert Kessler	**School of Medicine**
Dr. Frank Polansky	**School of Medicine**

Associated faculty:

Prof. Edison Tse	**Engineering-Economic Systems**
Prof. Ross Shachter	**Engineering-Economic Systems**

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The Computer-Aided Medical Decision Analysis (CAMDA) project is an attempt to develop intelligent medical decision systems by taking advantage of the complementary methodologies of decision analysis and artificial intelligence.

B. Medical Relevance and Collaboration

The primary effort of the CAMDA project during 1983 was focused on the design and implementation of RACHEL, an intelligent decision system for infertile couples. This effort is aimed at helping physicians and patients deal with difficult choices regarding pertinent medical procedures. RACHEL is being developed in close cooperation between the department of Engineering-Economic Systems, the department of Obstetrics and Gynecology, and the department of Surgery (Urology Division), all at Stanford.

C. Highlights of Research Progress

C.1 Accomplishments this past year

The CAMDA project began in the summer concentrated our efforts on three specific tasks: the development of a formal representation for uncertain decisions, the design and implementation of solution algorithms for formal decision problems, and the construction of an inferential processor specifically tailored to the process of formalizing decision problems.

Most of our research has been based on the concept of an influence diagram (Howard and Matheson, 1984) which is generalization of decision trees as a representation for decision problems. Influence diagrams (IDs) have several major features that make them attractive for use in intelligent decision systems. Technically, IDs prevent the loss of information often incurred in constructing asymmetric decision trees (Olmsted, 1984),

without suffering from the explicit exponential growth of symmetric decision trees. Furthermore, unlike decision trees, ID decision models take full advantage of probabilistic independence relations, which can have a significant impact on the simplicity of the decision model, and on the efficiency of its formal solution.

A particularly useful feature of influence diagrams which we have recently begun investigating is that fact that they can be used to represent deterministic as well as probabilistic relations between model elements. In fact, deterministic relations can be exploited to describe complex probabilistic behavior. This feature allows the construction of simple knowledge bases (composed primarily of deterministic statements) which can be used to create problem-specific probabilistic decision models.

In addition to their technical advantages IDs have been empirically shown to be intuitively appealing to decision makers (Owen, 1978), and to provide an excellent means of communication between experts in different fields. In the context of our own efforts, we have found that the physicians who are participating in the development of RACHEL have had little difficulty using IDs as a simple representation for expressing the decisions they and their patients face.

Another important feature of IDs is the fact that they are naturally constructed in a backwards, goal-directed fashion (decision trees usually lead to a forward-reasoning approach). Backward development of decision models has two important advantages for our purposes. First, it has a strong attention-focusing effect since it encourages the decision maker to first think of what he or she wants, and then about what can be done to change the world according to the expressed preferences. Decision trees usually have the opposite effect. Thus, they often lead the decision process along paths that although relevant to the decision at hand, have little effect on it. The attention-focusing effect of IDs on the decision making process tends to contribute to its efficiency. The second advantage of the goal-directed nature of IDs for the construction of intelligent decision systems is that it makes the formulation of decision problems amenable to computer-based automation as a rule-based system.

Having decided on IDs as a means to represent decision problems, we have designed and implemented several algorithms to solve well-formed influence diagrams. This effort has resulted in the development of a powerful software package which can generate optimal strategies and their certain equivalent directly from an ID. This package is beginning to be tested and augmented to make it easier to use by researchers other than its developers. In part, this package is based on the work of Olmsted (1983), and on a constructive proof by Shachter (1984) that, given certain technical features, shows that an influence diagram can always be solved in finite time.

An important feature of RACHEL is that it attempts to help its users in the development of models for their decisions. Thus, unlike most other decision analysis tools, RACHEL is designed to use domain knowledge. Therefore, a central element in the architecture of the RACHEL system is an algorithm which performs symbolic inference. Although several general-purpose inference-engines exist within our research environment, we have found it advantageous to implement our own for reasons of efficiency and compatibility. Furthermore, our inference algorithms are particularly well suited for the construction of decision-analytic models.

Finally, from the standpoint of computer implementation, we have developed a data structure which allows us to represent a wide class of multiple-entry disconnected cyclical directed graphs, where both vertices and edges can be associated with arbitrary data structures (such as frames). For short, we refer to these graphs as WEBs (as in a spider's), and we have used them to represent a multitude of small and medium-sized

objects such as influence diagrams, medical decision knowledge bases, command parse tables, help text databases, and mathematical data (e.g., vectors and matrices).

C.2 Research in progress

The immediate goal of the CAMDA project is to complete a pilot-level implementation of the RACHEL system within the next few months. As we define it, a pilot system is one where the essential algorithms work both individually and interactively with one another, operating with knowledge that is representative of the system's domain. Such a system lacks two important elements that must exist within a prototype-level implementation: an extensive knowledge base, and a front end usable by trained users who may not be familiar with the details of the system.

To complete a pilot implementation of RACHEL, we intend to direct our efforts towards the following four tasks: incorporating a medical value model elicitation facility, strengthening our influence-diagram solution procedure, improving the performance of RACHEL's inference engine, and implementing an explanation module to justify the decision model being developed. Once this implementation is completed, RACHEL will be brought to the participating physicians to begin to develop its knowledge base.

D. Publications

1. Breese, J.S., Davis, D., Parnell, G.S., and Taneja, R.: *"IDEAS: Influence Diagram Elicitation System"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

2. Holtzman, S.: *"A Model of the Decision Analysis Process"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1981.

3. Holtzman, S.: *"A Decision Aid for Patients with End-Stage Renal Disease"*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

4. Holtzman, S.: *"On the Use of Formal Models in Decision Making"*, Proc. TIMS/ORSA Joint Nat. Mtg., San Francisco, May, 1984.

5. (*) Holtzman, S.: *"Intelligent Decision Systems"*, Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, forthcoming.

6. Howard, R.A., and Matheson, J.E.: *"Influence Diagrams"*, in Howard, R.A., and Matheson, J.E. (Eds.): "The Principles and Applications of Decision Analysis," Vol. II, Strategic Decisions Group, Menlo Park, California, 1984.

7. Olmsted, S.M.: *"On Representing and Solving Decision Problems"*, Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

8. Owen, D.L.: *"The Use of Influence Diagrams in Structuring Complex Decision Problems"*, in Howard, R.A., and Matheson, J.E. (Eds.): "The Principles and Applications of Decision Analysis," Vol. II, Strategic Decisions Group, Menlo Park, California, 1984.

9. Shachter, R.: *"Evaluating Influence Diagrams"*, working paper, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1984.

E. Funding Support

The CAMDA project does not yet have direct funding support. However, in addition to SUMEX computer usage, the project has benefited from a number of hardware gifts and research support for individuals.

E.1 Stanford Medical School

The department of Obstetrics and Gynecology and the department of Surgery (Urology Division) have provided various types of support to the project. Samuel Holtzman has received research assistantship awards for several quarters. In addition, the Infertility Clinic at Stanford has purchased several terminals for the specific purpose of developing RACHEL and other CAMDA decision systems.

E.2 Decision Systems Laboratory

The CAMDA project has access to the facilities of the Decision Systems Laboratory (DSL) in the Department of Engineering-Economic Systems, and constitutes the laboratory's most active research project. The DSL maintains several terminals, printers and a personal computer for research on the development of computer-based decision systems. The majority of the terminals and printers were recently donated to the DSL by Qume Corporation. MAD Computer of Santa Clara has also contributed to the support of the CAMDA project through the consignment of a MAD-1 personal computer, and provision of a research assistantship for Samuel Holtzman.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

II.A Medical Collaborations and Program Dissemination Via SUMEX

Since its inception, the CAMDA project has benefited from an active relationship between decision analysts, computer scientists, and members of the Stanford medical community. In particular, RACHEL is being developed in close cooperation with physicians in the Infertility Clinic at Stanford. Most of this cooperation has, up to this point, consisted of an intense mutual learning experience for all project participants. The primary purpose of this initial effort has been to develop an effective means to represent medical decision knowledge. As we have described above, this work has culminated in the definition of a representation language based on influence diagrams.

Within the next few months, RACHEL is expected to attain pilot-level performance, and its knowledge base will begin to be developed. At this point, most of the interaction involving participating physicians will shift to the design and implementation of an infertility decision knowledge base. This task will involve considerable direct use of the SUMEX facility by medical personnel.

As an added benefit of the development of RACHEL, it often occurs that specific subsystems become useful in their own right. For instance, a simple program to aid physicians in determining a course of action in cases of idiopathic infertility has been implemented and made available on SUMEX to the staff of the Stanford infertility clinic, who have used it on an experimental basis.

II.B. Sharing and Interactions with other SUMEX-AIM Projects

II.B.1 SUMEX-AIM 1983 Workshop:

Samuel Holtzman chaired the working group on decision analysis and artificial intelligence in medicine. This group considered the current status and future of medical

decision systems. A full report of the working group's deliberations and conclusions is available online at SUMEX in file <HOLTZMAN.CAMDA>AIM-DA-FINAL-REPORT.TXT, and should appear in the forthcoming workshop report.

II.B.2 Participation in the Knowledge Representation seminar at Stanford

As part of the CAMDA project, we have made several presentations to the general Stanford medical and computer science community. These presentations have been made within the context of the Knowledge Representation seminar, held jointly by the computer science department and the medical school, and well attended by other SUMEX researchers at Stanford.

The speakers, and titles of the most recent presentations follow:

Samuel Holtzman:	On the Design and Implementation of Computer-Based Decision Systems.
Samuel Holtzman:	A Simple Representation for Uncertain Knowledge
Prof. Ross Shachter:	Influence Diagrams and their Use in Representing and Solving Complex Decision Problems.

II.C. Critique of Resource Management

The CAMDA project has been immeasurably aided by the availability of the SUMEX computing resources. In general we find the overall physical facilities to be of excellent quality. In addition, we have been quite impressed with the quality of the SUMEX staff. In particular, we have found it to be a pleasure to deal with Ed Pattermann, who has been invariably courteous, responsive to our needs, and effective in his actions. Pam Ryalls has also provided much needed help in managing the CAMDA project in a manner that is friendly and efficient.

There are, however, two areas where we feel service and performance could be improved to the benefit of the entire SUMEX community. The first concerns the SUMEX facilities themselves, the other refers to our means of communicating with these facilities.

II.C.1 SUMEX load

In the period that the CAMDA project has been active, we have noticed a significant increase in the maximum machine loading, particularly during weekday afternoons. Although this is a normal feature of time-shared computer systems, the load has become sufficiently high in recent months, that it is beginning to be difficult to work on SUMEX during business hours. In addition, reliability has been adversely affected in some instances. An increase in SUMEX computing capacity, or a means of preventing overloading of the machine should be considered. We believe that an emphasis on distributing some of SUMEX's computing power away from a centralized mainframe could have a significant effect on reducing the system load.

II.C.2 Ethernet

The CAMDA project uses SUMEX almost exclusively through Ethernet software and hardware located in Terman Engineering Center, where the department of Engineering-Economic Systems is located. This software has on occasion been extremely unreliable for extended periods of time, resulting in substantially reduced productivity for the project. Adequate communication facilities at Stanford are of critical importance to

the successful conduct of our research. Although Stanford Ethernet management is not directly under the jurisdiction of SUMEX, in order for the SUMEX resource to be utilized to the fullest, the planning and administration of networking at Stanford needs to be better coordinated. We have begun to explore several means to improve the current situation, and we believe that explicit SUMEX support of our efforts would be quite beneficial.

III. RESEARCH PLANS

III.A Project Goals and Plans

For the near term future the primary goals of the CAMDA project are to develop a "pilot" and then "prototype" version of the RACHEL system. Over an extended period, our objective is to arrive at useable, fully-validated and documented systems for support of medical decision making in infertility and other domains.

Implementation of a pilot system is primarily an integrative task at this point, bringing together the medical knowledge base, symbolic inference procedure, decision problem solution procedure, and influence diagram data structures. All of these components exist independently. The pilot system will consist of these systems interacting to provide a simplified version of infertility decision counseling.

The prototype implementation of RACHEL will include substantially greater amounts of medical knowledge than the pilot. The major task at this stage will be the incorporation of expert knowledge regarding functional relations, probability distributions, and decision alternatives in the infertility domain. At this point in its development, the system will be available for use by participating physicians at the infertility clinic on a "test" basis, beginning the critical phase of validation and justification of the system.

A major goal of the project is to bring RACHEL to a "defensible" level of performance in the infertility domain. A working system with full documentation, explanation of its conclusions, and user interface is envisioned. Over the long term, infertility is but a single example of the range of medical decisions amenable to decision analytic treatment in an automated system. After RACHEL has been fully implemented and tested, other systems focusing on cardiology or oncology, for example, might be developed. These systems would consist of a common core of procedural knowledge based on decision analysis, and be instantiated with the medical knowledge of the particular domain.

III.B Justification and Requirements for Continued SUMEX Use

The CAMDA project is truly interdisciplinary. It draws on elements of decision analysis, artificial intelligence, and medical science. The project has the potential to contribute to each of these disciplines in important ways. SUMEX-AIM provides the resources to continue this research with the necessary access to members of the Stanford research community.

The development of automated decision systems has the potential to greatly increase the use and acceptance of decision analysis methods. In the past, although decision analysis has been shown to be an extremely effective means of assisting in decision making in complex and uncertain domains, the cost and effort involved in producing an analysis was prohibitive for most individuals. Automated decision analysis can result in a much lower cost per user, allowing decision theoretic techniques to achieve much wider application.

The development of decision systems owes much to advancements in the fields of

artificial intelligence, expert systems, and knowledge engineering. One continuing challenge in these fields has been representation and reasoning with probabilistic knowledge. The representation of knowledge in influence diagrams, and the use of decision analysis in probabilistic reasoning are both significant topics of research being pursued within the CAMDA project.

For the medical community the CAMDA project has the potential for providing tools and techniques that greatly improve the quality of decision making in medicine. RACHEL explicitly considers uncertainty, decision alternatives, and patient preferences in developing recommendations. The objective is to develop insight and understanding regarding tradeoffs and alternatives, both for the patient and the attending physician.

SUMEX-AIM provides a unique resource for the continuation of the CAMDA project. The available computing resources, plus access to the Stanford AI and medical communities are of critical importance for the successful completion of the research.

III.C Needs and Plans for other Computing Resources beyond SUMEX-AIM

We are pursuing the purchase or donation of several computing resources for installation in the Decisions System Laboratory. Our primary need at present is for a LISP machine (e.g., Symbolics 3600), enabling us to perform local processing and increase our graphics capabilities.

At present the project has access to one MAD-1 personal computer (IBM-PC type). We are considering various other PC/workstation facilities to use as front ends for CAMDA products.

III.D Recommendations for Future Community and Resource Development

Increases in distributed computing capabilities on the SUMEX-AIM system is a primary need at this point. As we mentioned in Section II.C.1, distributed file editing and graphics capabilities would simultaneously reduce load on the mainframe. At this time, we are particularly interested in the possibility of designing an environment where a centralized processor (such as the SUMEX 20/60) would interact at a high level with a much less powerful dedicated processor (such as a SUN workstation, or an Apple Lisa or Macintosh) with specific capabilities such as bit-mapped graphics and special purpose hardware.

II.A.3.2. MENTOR Project

MENTOR Project

Stuart M. Speedie, Ph.D.
Terrence F. Blaschke, M.D.
Department of Medicine
Division of Clinical Pharmacology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goal of the MENTOR (Medical EvaluatioN of Therapeutic ORders) project is to design and develop an expert system for monitoring drug therapy for hospitalized patients that will provide appropriate advice to physicians concerning the existence and management of adverse drug reactions. The computer as a recording-keeping device is becoming increasingly common in hospital-based health care, but much of its potential remains unrealized. Furthermore, this information is provided to the physician in the form of raw data which is often difficult to interpret. The wealth of raw data may effectively hide important information about the patient from the physician. This is particularly true with respect to adverse reactions to drugs which can only be detected by simultaneous examinations of several different types of data including drug data, laboratory tests and clinical signs.

In order to detect and appropriately manage adverse drug reactions, sophisticated medical knowledge and problem solving is required. Expert systems offer the possibility of embedding this expertise in a computer system. Such a system could automatically gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base of relevant data, it could analyze incoming data and inform physicians when adverse reactions are likely to occur or when they have occurred. The MENTOR project is an attempt to explore the problems associated with the development and implementation of such a system and to implement a prototype of a drug monitoring system in a hospital setting.

B. Medical Relevance and Collaboration

A number of independent studies have confirmed that the incidence of adverse reactions to drugs in hospitalized patients is significant and that they are for the most part preventable. Moreover, such statistics do not include instances of suboptimal drug therapy which may result in increased costs, extended length-of-stay, or ineffective therapy. Data in these areas are sparse, though medical care evaluations carried out as part of hospital quality assurance programs suggest that suboptimal therapy is common.

Other computer systems have been developed to influence physician decision making by monitoring patient data and providing feedback. However, most of these systems suffer from a significant structural shortcoming. This shortcoming involves the evaluation rules that are used to generate feedback. In all cases, these criteria consist of discrete, independent rules. Yet, medical decision making is a complex process in which many factors are interrelated. Thus attempting to represent medical decision-making as a

discrete set of independent rules, no matter how complex, is a task that can, at best, result in a first order approximation of the process. This places an inherent limitation on the quality of feedback that can be provided. As a consequence it is extremely difficult to develop feedback that explicitly takes into account all information available on the patient. One might speculate that the lack of widespread acceptance of such systems may be due to the fact that their recommendations are often rejected by physicians. These systems must be made more valid if they are to enjoy widespread acceptance among physicians.

The proposed MENTOR system is designed to address the significant problem of adverse drug reactions by means of a computer-based monitoring and feedback system to influence physician decision-making. It will employ principles of artificial intelligence to create a more valid system for evaluating therapeutic decision-making.

The work in the MENTOR project is intended to be a collaboration between Dr. Blaschke at Stanford and Dr. Speedie at the University of Maryland. Dr. Speedie is spending the 1983-84 academic year on sabbatical with Dr. Blaschke in the Division of Clinical Pharmacology at Stanford University. While at Stanford, Dr. Speedie has been strengthening his expertise in the area of artificial intelligence and establishing links in the AI community. Dr. Speedie has begun work on the development of the MENTOR system pilot project on the SUMEX-AIM facility. Over the past nine months, Drs. Blaschke and Speedie have worked closely together to design the MENTOR project. The blend of previous experience, medical knowledge, computer science knowledge and evaluation design expertise they represent is vital to the successful completion of the activities in the MENTOR project.

C. Highlights of Research Progress

The MENTOR project was initiated in December 1983. The work to date has consisted of preparation of a grant proposal for the National Center for Health Services Research and initial exploration of the problem of designing the MENTOR system. Work has begun on constructing a system for monitoring potassium in patients with drug therapy that can adversely affect potassium levels.

E. Funding Support

Application for grant support is pending.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

This project represents a collaboration between faculty at Stanford University Medical Center and the University of Maryland School of Pharmacy in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, will facilitate this collaboration when Dr. Speedie returns to the University of Maryland in August of 1984.

B. Sharing and Interactions with Other SUMEX-AIM Projects

Interactions with other SUMEX-AIM projects has been on an informal basis. Personal contacts have been made with individuals working on the ONCOCIN project concerning issues related to the formulation of the previously mentioned proposal. We expect interactions with other projects to increase significantly once the groundwork has

been laid and issues directly related to AI are being addressed. Given the geographic separation of the investigators, the ability to exchange mail and programs via the SUMEX system as well as communicate with other SUMEX-AIM projects is vital to the success of the project.

C. Critique of Resource Management

To date, the resources of SUMEX have been fully adequate for the needs of this project. The staff have been most helpful with any problems we have had and we are fully satisfied with the current resource management. The only concern we have relates to the state of the documentation on the system.

III. RESEARCH PLANS

A. Project Goals and Plans

To accomplish the goals described in the Project Rationale, a number of tasks will be undertaken. The short-term task is to develop an initial prototype of the medical knowledge base and inference mechanisms for arriving at appropriate therapy monitoring decisions. This initial work focuses on monitoring for hyperkalemia and the decision-making process with respect to ordering potassium levels. We will then attempt to construct a system combining frames and rules that will model this process. The purpose of this initial exercise is to explore the problems involved in constructing an AI system that meets the needs of drug therapy monitoring and to establish development guidelines for the larger project.

The long-range plans for the MENTOR project depend on the outcome of the funding decision. However, assuming a favorable decision, the full project has the following goals:

1. Implement a prototype computer system to continuously monitor patient drug therapy in a hospital setting. This will be an expert system that will use a modular, frame-oriented form of medical knowledge, a separate inference engine for applying the knowledge to specific situations and automated collection of data from hospital information systems to produce therapeutic advisories.

2. Select a small number of important and frequently occurring medical settings (e.g., combination therapy with cardiac glycosides and diuretics) that can lead to therapeutic misadventures, construct a comprehensive medical knowledge base necessary to detect these situations using the information typically found in a computerized hospital information system and generate timely advisories intended to alter behavior and avoid preventable drug reactions.

3. Select and test several methods of formulating and providing advisories to physicians in order to find an optimal method of feedback that is acceptable and useful to physicians and is feasible to implement.

4. Design and begin to implement an evaluation of the impact of the prototype MENTOR system on physicians' therapeutic decision-making as well as on outcome measures related to patient health and costs of care.

B. Justification and Requirements for Continued SUMEX Use

This project needs continued use of the SUMEX facilities for two reasons. First is that it provides access to an environment specifically designed for the development of AI

systems. The MENTOR project focuses on the development of such as system for drug monitoring that will explore some neglected aspects of AI in medicine. Access to SUMEX is necessary for timely development of the MENTOR system, as well as advice and assistance in the design and development of a well-designed and efficient system. Access to SUMEX is also necessary to support the collaborative effort in this project as described previously.

C. Needs and Plans for Other Computing Resources Beyond SUMEX-AIM

A major long-range goal of the MENTOR project is to implement this system on a independent hardware system of suitable architecture. It is recognized that the full monitoring system will require a large patient data base as well as a sizeable medical knowledge base and must operate on a close to real-time basis. Ultimately, the SUMEX facilities will not be suitable for these applications. Thus we intend to transport the prototype system to a dedicated hardware system that can fully support the the planned system and which can be integrated into the SUMC Hospital Information System. However, no firm decisions have been made about the requirements for this system since many specification and design decisions remain to be made.

D. Recommendations for Future Community and Resource Development

In the brief time we have been associated with SUMEX, we have been generally pleased with the facilities and services. However, it is evident that disk space is a critical factor in the functioning of the facility. It would seem wise to increase disk storage in order to meet the needs of the users. Our experience also indicates that an attempt needs to be made to organize and update the documentation associated with the various SUMEX systems. Being new users, we found that paths to useful software was somewhat longer than one might expect. An expanded introduction to the system that, at least, briefly described the software available on SUMEX would be useful.

II.A.3.3. Protein Secondary Structure Project

Protein Secondary Structure Project

Robert M. Abarbanel, M.D.
Section on Medical Information Science
University of California Medical Center
University of California at San Francisco

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Development of a protein structure knowledge base and tools for manipulation of that knowledge to aid in the investigation of new structures. System to include cooperating knowledge sources that work under the guidance of other system drivers to find solutions to protein structure problems. Evaluations of structure predictions using known proteins and other user feedbacks available to aid user in developing new methods of prediction.

B. Medical Relevance and Collaboration

Many important proteins have been sequenced but have not, as yet, had their secondary or tertiary structures revealed. The systems developed here would aid medical scientists in the search for particular configurations, for example, around the active sites in enzymes. Predictions of secondary structure will aid in the determination of the full "natural" configuration of important biological materials. Development of systems such as these will contribute to our knowledge of medical scientific data representation and retrieval.

C. Highlights of Research Progress

The prediction of beta-alpha protein structures is complete. The system was developed on a VAX 11/750 at the University of California, San Francisco, to allow researchers to describe patterns of amino acid residues that will be sought in the sequences under study. The presence or absence of these "primary" patterns are then combined with other measures of structure, like hydrophobicity, to suggest possible alpha helix or beta sheet or turn configurations.

The segments of a sequence between turns are then analyzed to determine the allowable extent of the possible secondary structure assignments. Any segments remaining are then used to generate all possible complete structures. Only two beta strands with the character of sheet edges are allowed in any prediction. This hierarchical generation and pruning results in nearly 95% turn prediction accuracy, and excellent delimiting of helices and sheets. In some cases, one and only one secondary structure is predicted.

Research in Progress -- At this time, work is under way to extend this α/β assignment work to a set of cancer causing viral proteases. These proteins are believed to be of the α/β type. The set of homologous sequences under study introduces new problems and insight into the problems of structural assignment. If one is to believe that major structural features are conserved across a primary sequence homology, then methods must be developed for predicting structure when possibly conflicting signals come

from individual sequences in a set. Other sets of proteins, like the Triose Phosphate Isomerases, will help to develop this knowledge.

Dr. F. Cohen is using the pattern matching and rules system on a regular basis to develop a means for predicting turns in proteins of the all-α and all-β classes. His use of the system stimulates simultaneous development of improved rules support and explanation facilities.

D. List of Relevant Publications

The first paper on $\beta\alpha\beta$ protein structures has been published: Cohen, F.E., Abarbanel, R.M., Kuntz, I.D. and Fletterick, R.J.: ***Secondary structure assignment for α/β proteins by a combinatorial approach***, Biochemistry, 22, pp 4894-4909, (October 1983). At this time, another paper on prediction of "turns" in several classes of proteins is under preparation. Similar pattern matching tools are implemented in the QUEST program written in Mainsail and supported commercially by Intelligenetics, Inc. This program converts patterns given by users into and/or trees of finite state machines: Abarbanel, R.M., P.R. Wieneke, E. Mansfield, D.A. Jaffe, and D.L. Brutlag, ***Rapid searches for complex patterns in biological molecules***, Nucleic Acids Research, 12, pp 263-280, (January 1984).

E. Funding Support

Title:	Protein Structural Knowledge Engineering
Principal Investigator:	Robert M. Abarbanel, M.D.
Funding Agency:	National Library of Medicine, N. I. H.
Grant ID Number:	1 R23 LM 03893-01
Total Award:	4/1/83 to 3/31/86
	$ 104876 Total Direct Costs
Current Period:	4/1/84 to 3/31/85
	$ 40900 Total Direct Costs

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

None.

B. Sharing and Interactions with SUMEX Projects

This project is closely allied with the MOLGEN group, both in computer and scientific interests. Some pattern matching methodology created for the protein data base has been adopted and used in the various DNA knowledge bases. The principal persons in the MOLGEN group have contributed to this project's use and understanding of knowledge base software and resources.

C. Critique of Resource Management

Work continues on the UNIX systems at the University of California, San Francisco. SUMEX has been used primarily for communications with other researchers. At some future date it is expected that the knowledge based system will be ported to SUMEX on one or more of the LISP machines available.

Resource management remains excellent. The staff are friendly and responsive.

Network access, bulletin boards and the mail system have provided a means to collaborate with others doing related work locally as well as in Europe. SUMEX-AIM staff have been most helpful in getting this project started on the Dolphin workstations and in providing an environment where new tools have been made available for use.

III. RESEARCH PLANS

A. Project Goals and Plans

Near Term -- Development of "parallel" assignment techniques to allow homologous sequences to aid in the prediction of structure for one or more unknown sequences. Completion of Lisp system providing a friendly environment for structure exploration. This will involve merging sequential rule interpretation with back chaining. Both these systems will be able to invoke the running of patterns of amino-acid residues against known or unknown sequences. Along with the capacity to manipulate the order of application of rules, the system will allow undoing of decisions during processing, and explanation of reasoning during structure assignment. These are all features of knowledge engineering that are not present in the current system.

Long Term -- Expansion of techniques used for α/β prediction to other classes of proteins. Improvement of user interfaces to allow use of this sequence analysis system for problems of homology and energetics. Use of bit-map graphics and an interface to the line-drawing color graphics at UCSF to enhance the user's view of the data and possibly enhance the development of new knowledge sources for application to these problems. Several areas of current interest may contribute here: distance geometry, docking, energy minimization, and multi-sequence homologies.

B. Need for Resources

SUMEX Resources -- The availability of UNIX (TM) under SUMEX-AIM control will greatly aid in the transferability of existing algorithms. The environment of knowledge base tools and people is the primary motive for doing this work using SUMEX. Access to both established and developing systems aids this project in setting down standards of excellence, forward thinking about computing tools and methodologies, and active exchange of techniques and ideas. The close collaboration with the MOLGEN researchers is particularly useful in this regard.

Other Computing Resources -- A soon to be established network connection with the Computer Graphics Laboratory at UCSF will provide access to 1) the latest in protein structural information, and 2) color line drawing graphics facilities for evaluation and display of this projects product. A real time display using color graphics will become a possibility. Lisp based machines soon to be acquired at UCSF will allow direct collaboration with efforts at SUMEX on knowledge based software for protein structure determination.

C. Recommendations

No changes from last year's report: First and most important -- EXPAND the computing power available to SUMEX users. Facilitate networking with other computing environments like the Computer Graphics Laboratory at UCSF so that protein structural information may be exchanged and their hardware for 3D structure display may be utilized as a part of a complete biological structures analysis system.

Second -- Provide whatever hand-holding is necessary to expose SUMEX-AIM users

to other facilities available on the network. This will allow a project to find its best home in the SUMEX environment.

II.A.3.4. PROTEAN Project

PROTEAN Project

Oleg Jardetzky
Nuclear Magnetic Resonance Lab, School of Medicine
Stanford University

Bruce Buchanan
Computer Science Department
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale The goal of this project is two-fold: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallizing proteins), and (b) use protein structure determination as a test problem for experiments with the AI control structure known as the Blackboard Model.

B. Medical Relevance The molecular structure of proteins is essential for understanding many problems of medicine at the molecular level, such as the mechanisms of drug action. Using NMR data from proteins in solution will speed up the determination.

C. Highlights of Progress This project is just getting started. There is no substantial progress to date.

E. Funding Support

Grant applications submitted to the NSF:

Title: Interpretation of NMR Data from Proteins
Using AI Methods

PI's: Oleg Jardetzky and Bruce G. Buchanan

Agency: National Science Foundation

Total Amount: $969,991.

Dates: Apr 1, 1984/March 31, 1989

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations Several members of Prof. Jardetzky's research group are involved in this research.

B. Interactions with other SUMEX-AIM projects

Robert Langridge is visiting in the HPP this academic year and has been participating in discussions. Carroll Johnson has been helpful in making ORTEP available and answering questions about it.

C. Critique of Resource Management

The SUMEX staff has been most cooperative in helping get this project started. Because the terminals available in SMRL for our use are IBM PC's, we needed considerable help with communications.

III. RESEARCH PLANS

A. Goals & Plans

Our long range goal is to build an automatic interpretation system similar to CRYSALIS(which worked with x-ray crystallography data). In the shorter term, we are building interactive programs that aid in the interpretation. We are putting together building blocks now and are designing the control structure. We plan to purchase a high resolution graphics display workstation as soon as our exploratory investigations indicate the expense is justified.

B. Justification for continued SUMEX use

We will continue to use SUMEX for developing the AI methods. We need Interlisp to implement the Blackboard model and knowledge structures most flexibly and quickly.

C. Need for other computing resources

We believe we must purchase a graphics workstation for display of partial results.

D. Recommendations

With the increased number of personal computers and workstations in the community, it would be desirable to provide more staff to integrate these machines with SUMEX and centralize sharing of software across the community.

II.A.3.5. Ultrasonic Imaging Project

Ultrasonic Imaging Project

James F. Brinkley, M.D.
W.D. McCallum, M.D.
Depts. Computer Science, Obstetrics and Gynecology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The long range goal of this project is the development of an ultrasonic imaging and display system for three-dimensional modelling of body organs. The models will be used for non-invasive study of anatomic structure and shape as well as for calculation of accurate organ volumes for use in clinical diagnosis. Initially, the system has been used to determine fetal volume as an indicator of fetal weight; later it will be adapted to measure left ventricular volume, or liver and kidney volume.

The general method we are using is the reconstruction of an organ from a series of ultrasonic cross-sections taken in an arbitrary fashion. A real-time ultrasonic scanner is coupled to a three-dimensional acoustic position locating system so that the three- dimensional orientation of the scan plane is known at all times. During the patient exam a dedicated microcomputer based data acquisition system is used to record a series of scans over the organ being modelled. The scans are recorded on a video tape recorder before being transferred to a video disk. 3D position information is stored on a floppy disk file. In the proposed system the microprocessor will then be connected to SUMEX where it will become a slave to an AI program running on SUMEX. The SUMEX program will use a model appropriate for the organ which will form the basis of an initial hypothesis about the shape of the organ. This hypothesis will be refined at first by asking the user relevant clinical questions such as (for the fetus) the gestational age, the lie of the fetus in the abdomen and complicating medical factors. This kind of information is the same as that used by the clinician before he even places the scan head on the patient. The model will then be used to request those scans from the video disk which have the best chance of giving useful information. Heuristics based on the protocols used by clinicians during an exam will be incorporated since clinicians tend to collect scans in a manner which gives the most information about the organ. For each requested scan a two-dimensional tolerance region (or plan) derived from the model will be sent to the microcomputer. The requested scan will be retrieved from the video disk, digitized into a frame buffer, and the plan used to direct a border recognition process that will determine the organ outline on the scan. The resulting outline will be sent to SUMEX where it will be used to update the model. The scan requesting process will be continued until it is judged that enough information has been collected. The final model will then be used to determine volume and other quantitative parameters, and will be displayed in three dimensions.

We believe that this hypothesize verify method is similar to that used by clinicians when they perform an ultrasound exam. An initial model, based on clinical evidence and past experience, is present in the clinician's mind even before he begins the exam. During the exam this model is updated by collecting scans in a very specific manner which is known to provide the maximum amount of information. By building an ultrasound imaging system which closely resembles the way a physician thinks we hope to not only

provide a useful diagnostic tool but also to explore very fundamental questions about the way people see.

We are developing this system in phases, starting with an earlier version developed at the University of Washington. During the first phase the previous system was adapted and extended to run in the SUMEX environment. Clinical studies were done to determine its effectiveness in predicting fetal weight. In the second phase computer vision techniques were used to solve some of the problems observed in the clinical trials on the first phase. Further iterations will be tested against clinical data, thus providing valuable feedback for the development process.

B. Medical Relevance and Collaboration

This project is being developed in collaboration with the Ultrasound Division of the Department of Obstetrics at Stanford, of which W.D. McCallum is the director.

Fetal weight is known to be a strong indicator of fetal well-being: small babies generally do more poorly than larger ones. In addition, the rate of growth is an important indicator: fetuses which are "small-for-dates" tend to have higher morbidity and mortality. It is thought that these small-for-dates fetuses may be suffering from placental insufficiency, so that if the diagnosis could be made soon enough early delivery might prevent some of the complications. In addition such growth curves would aid in understanding the normal physiology of the fetus. Several attempts have been made to use ultrasound for predicting fetal weight since ultrasound is painless, noninvasive, and apparently risk-free. These techniques generally use one or two measurements such as abdominal circumference or biparietal diameter in a multiple regression against weight. We recently studied several of these methods and concluded that the most accurate were about +/-200 gms/kg, which is not accurate enough for adequate growth curves (the fetus grows about 200 gms/week). The method we have developed is based on the fact that fetal weight is directly related to volume since the density of fetal tissue is nearly constant. We showed last year that by utilizing three dimensional information more accurate volumes and hence weights can be obtained.

In addition to fetal weight, the first implementation of this system has been evaluated for its ability to determine other organ volumes in vitro. In collaboration with Dr. Richard Popp of the Stanford Division of Cardiology we have evaluated the system on in vitro kidneys and latex molds of the human left ventricle. Left ventricular volumes are routinely obtained by means of cardiac catheterization in order to help characterize left ventricular function. Attempts to determine ventricular volume using one or two dimensional information from ultrasound has not demonstrated the accuracy of angiography. Therefore, three-dimensional information should provide a more accurate means of non-invasively assessing the state of the left ventricle.

C. Highlights of Research Progress

In the last report an initial version of the second phase of program development was described. This version utilizes AI techniques to solve some of the problems encountered with the non-AI system. The prototype system was implemented and tested on two shape classes of balloons (round and long-thin).

For each balloon class a training set of similarly-shaped balloons was used to give the computer knowledge of the given shape. This training set consisted of ultrasonic reconstructions obtained by the previous system. The knowledge was then used to analyze ultrasound data from a similarly-shaped balloon which was not part of the training set. The initial input to the system consisted of the three-dimensional positions

and orientations of a series of ultrasound slices. These slices were previously acquired manually and stored on a video tape recorder. The system was also given the two endpoints of the balloons, which allowed a reference coordinate system to be established. The balloon endpoints interacted with the shape knowledge to define an initial tolerance region, within which the system expected the actual balloon surface to be found. The system's best guess as to the location of the actual balloon surface was the middle of the tolerance region.

Once the initial tolerance region was established an hypothesize-verify paradigm was employed to alternately request a particular ultrasound slice, to provide a tolerance region for an edge detector on that slice, to manually acquire the border of the balloon on that slice, and to update the model by combining the new data with the shape knowledge. This process continued until it was judged that additional slices could contribute no new information.

For an example round balloon (measured volume 267 cc) the initial best guess volume after specifying the endpoints was 242 cc. After one slice best guess volume was 279 cc. After nine slices (out of a possible 30) the system judged that no more slices would be useful: best guess volume was 265 cc. For a different training set of long-thin balloons the final best guess volume for a new reconstruction, after 9 out of a possible 22 slices, was 459 cc, measured volume 461 cc. These results show that learned shape knowledge allowed the system to form a reasonable guess as to the location of the balloon surface even after only two endpoints had been specified.

The major accomplishment this past year was the compilation of the results from this project into the Ph.D. thesis of James Brinkley. In addition the artificial intelligence portion of the system was presented at several meetings, including the student paper competition of the Symposium on Computer Applications in Medicine, where it received the second place award.

Current research is suspended until I find a position following the Ph.D. There is currently some possibility of continuing the research on SUMEX at Stanford.

D. Recent Publications

1. Brinkley, J.F., Muramatsu, S.K., McCallum, W.D. and Popp, R.L.: ***In vitro evaluation of an ultrasonic three-dimensional imaging and volume system.*** Ultrasonic Imaging, 4:126-139, 1982.

2. Brinkley, J.F., McCallum, W.D., Muramatsu, S.K. and Liu, D.Y.: ***Fetal weight estimation from ultrasonic three-dimensional head and trunk reconstructions: Evaluation in vitro.*** Amer. J. Obstet. Gynecol. 144(6):715-721, 1982.

3. Brinkley, J.F., McCallum, W.D., Muramatsu, S.K., and Liu, D.Y.: ***Fetal weight estimation from lengths and volumes found by ultrasonic three-dimensional measurements.*** To be published in *Journal of Ultrasound in Medicine.*

4. Brinkley, J.F.: ***Artificial intelligence and ultrasonic imaging: the use of learned shape knowledge to analyze 3D data.*** Proceedings, 28th Annual Meeting, American Institute of Ultrasound in Medicine, New York, October, 1983.

5. Brinkley, J.F.: ***Learned shape knowledge in ultrasonic three-dimensional organ modelling.*** Second place, student paper competition, Symposium on Computer Applications in Medical Care, Baltimore, October 23-26, 1983.

6. Brinkley, J.F.: *Ultrasonic three-dimensional organ modelling.* Ph.D. Dissertation. Stanford University, to be published as a Stanford Computer Science Technical Report, Spring 1984.

7. Brinkley, J.F.: *Knowledge-driven ultrasonic three-dimensional organ modelling.* Submitted to *IEEE Trans. Pattern Analysis and Machine Intelligence.*

E. Funding Support

"Ultrasonic Three-dimensional Organ Modelling", individual postdoctoral fellowship. Fellow: James F. Brinkley Sponsor: W.D. McCallum Funding Agency: National Institute of General Medical Sciences Number: 1 F32 GM08092 Total term and direct cost: 7/1/81-6/30/84 (3 years) $55,452 (stipend) Current funding from this fellowship: 7/1/83-6/30/84 (1 year) $19,716

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

We are collaborating more with medical people than anyone else. The project is located in the Obstetrics Department at Stanford where W.D. McCallum manages the ultrasound patients. We have also been collaborating with Dr. Richard Popp in the Division of Cardiology at Stanford.

B. Sharing and Interactions with SUMEX projects

Mostly personal contacts with the Heuristic Programming Project and Medical Information Science Program at Stanford. The message facilities of SUMEX have been especially useful for maintaining these contacts. Since the first phase of the project is now essentially complete we have been interacting more with other SUMEX projects in order to develop the AI ideas.

C. Critique of Resource Management

In general SUMEX has been a very usable system, and the staff has been very helpful.

III. RESEARCH PLANS

A. Project Goals and Plans

The major conclusion from the research leading to the Ph.D. is that the current hardware we use for three-dimensional location is not accurate enough to permit further work on organ modelling. For this reason I have proposed several alternative methods of utilizing 3D medical image data, including 3D CT, NMR or ultrasound. All these modalities produce 3D arrays of data which would be much easier to use than arbitrary slices.

Given this type of data, fairly straightforward extensions of the model representation developed for balloons could be used for the heart or kidney. The basic idea would be to have the human operator indicate three organ landmarks within the 3D data, then let the computer utilize learned shape knowledge to selectively "biopsy" portions of the 3D data in order to define the actual organ instance. Since the data would be available as a 3D array, the edge detection process could take place along a one-

dimensional tolerance region rather than on a two-dimensional slice. Since all forms of medical images are becoming available as 3D arrays this seems like a better approach than the selection of individual slices.

Depending on the interest of engineers in providing 3D data much of the AI modelling could still be done on SUMEX. Many of the AI techniques could also be developed for 2D images for knowledge-driven border detection.

B. Justification and requirements for continued SUMEX use

The goals of this project seem to be compatible with the general goals of SUMEX, i.e., to develop the uses of artificial intelligence in medicine. The problem of three-dimensional modelling is a very general one which is probably at the heart of our ability to see. By developing a medical imaging system that models the way clinicians approach a patient we should not only develop a useful clinical tool but also explore some very fundamental problems in AI.

The availability of a large well supported facility like SUMEX has been and will continue to be very valuable as we develop and test further implementations of the system. Our current share of the SUMEX resources is adequate.

C. Needs and plans for other computing resources beyond SUMEX-AIM

Judging from our present experience it appears that SUMEX could not handle the amount of data required for image processing on digitized ultrasound scans. This is one of the main reasons we are proposing a distributed system in which SUMEX only directs a smaller machine to do the actual number crunching. It is also one of the reasons we are postponing direct digitization until later. As microprocessors become more powerful they will be capable of acting as slaves to an intelligent SUMEX program. The AI program will direct the image processing functions of the micro so that the data is processed in an intelligent way, but SUMEX will only see the results of that processing, not the actual data. We will thus need to keep track of developments in microcomputers so that we can develop this kind of distributed system.

An additional problem is the small address space of the 2060. Attempts will be made to optimize the code, but this could become a major problem in the future. A better solution might be an image processing workstation with a large address space.

D. Recommendations

Since we are planning to develop a distributed system we would hope to see these kind of systems being developed by the SUMEX resource. Projects that would be of direct interest are networks (such as ETHERNET), personal computer stations, graphics displays, etc.

II.A.4. Pilot AIM Projects

Following is a description of the informal pilot project currently using the AIM portion of the SUMEX-AIM resource, pending funding, full review, and authorization.

In addition to the progress report presented here, an abstract is submitted on a separate Scientific Subproject Form.

II.A.4.1. PATHFINDER Project

PATHFINDER Project

Bharat Nathwani, M.D.
Department of Anatomical Pathology
City of Hope National Medical Center
Duarte, California

Lawrence M. Fagan, M.D., Ph.D.
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Our project addresses difficulties in the diagnosis of lymph node pathology. Five studies from cooperative oncology groups have documented that, while experts show good agreement with one another, the diagnosis made by practicing pathologists may have to be changed by expert hematopathologists in as many as 50% of the cases. Precise diagnoses are crucial for the determination of optimal treatment. To make the knowledge and diagnostic reasoning capabilities of experts available to the practicing pathologist, we have developed a pilot computer-based diagnostic program called PATHFINDER. The project is a collaborative effort of the City of Hope National Medical Center and the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 45 common benign and malignant diseases of the lymph node based on 77 histologic features. Our research plans are to develop a full-scale version of the computer program by substantially increasing the quantity and quality of knowledge and to develop techniques for knowledge representation and manipulation appropriate to this application area. The design of the program has been strongly influenced by the INTERNIST/CADUCEUS program developed on the SUMEX resource.

A group of expert pathologists from several sites in the U.S., have agreed to help build the knowledge base for the PATHFINDER program. Each will independently provide the entire knowledge in incremental stages after agreement has been obtained on the design aspects. We estimate that the final version of the program will include about 80 diseases and 175 features.

B. Medical Relevance and Collaboration

One of the most difficult areas in surgical pathology is the microscopic interpretation of lymph node biopsies. Most pathologists have difficulty in accurately classifying lymphomas. Several cooperative oncology group studies have documented that while experts show good agreement with one another, the diagnosis rendered by a "local" pathologist may have to be changed by expert lymph node pathologists (expert hematopathologists) in as many as 50% of the cases.

The National Cancer Institute recognized this problem in 1968 and created the Lymphoma Task Force which is now identified as the Repository Center and the Pathology Panel for Lymphoma Clinical Studies. The main function of this expert panel

of pathologists is to confirm the diagnosis of the "local" pathologists and to ensure that the pathologic diagnosis is made uniform from one center to another so that the comparative results of clinical therapeutic trials on lymphoma patients are valid. An expert panel approach is only a partial answer to this problem. The panel is useful in only a small percentage (3%) of cases; the Pathology Panel annually reviews only 1,000 cases whereas more than 30,000 new cases of lymphomas are reported each year. A Panel approach to diagnosis is not practical and lymph node pathology cannot be routinely practiced in this manner.

We believe that practicing pathologists do not see enough case material to maintain a high-level of diagnostic accuracy. The disparity between the experience of expert hematopathology teams and those in community hospitals is striking. An experienced hematopathology team may review thousands of cases per year. In contrast, in a community hospital, an average of only 10 new cases of malignant lymphomas are diagnosed each year. Even in a university hospital, only approximately 100 new patients are diagnosed every year.

Because of the limited numbers of cases seen, pathologists may not be conversant with the differential diagnoses consistent with each of the histologic features of the lymph node; they may lack familiarity with the complete spectrum of the histologic findings associated with a wide range of diseases. In addition, pathologists may be unable to fully comprehend the conflicting concepts and terminology of the different classifications of non-Hodgkin's lymphomas, and may not be cognizant of the significance of the immunologic, cell kinetic, cytogenetic, and immunogenetic data associated with each of the subtypes of the non-Hodgkin's lymphomas.

In order to promote the accuracy of the knowledge base development we will have participants for multiple institutions collaborating on the project. Dr. Nathwani will be joined by experts from Stanford (Dr. Dorfman), St. Jude's Children's Research Center -- Memphis (Dr. Berard) and City of Hope (Dr. Burke).

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

Since the project's inception in November, 1983, we have constructed several versions of PATHFINDER. The first several versions of the program were ***rule-based*** systems like MYCIN and ONCOCIN which were developed earlier in the Stanford group. We soon discovered, however, that the large number of overlapping features in diseases of the lymph node would make a rule-based system cumbersome to implement. We next considered the construction of a ***hybrid system***, consisting of a rule-based algorithm that would pass control to an INTERNIST-like scoring algorithm if it could not confirm the existence of classical sets of features. We finally decided that a modified form of the INTERNIST program would be most appropriate. The current version of PATHFINDER is written in the computer language Maclisp and runs on the SUMEX DEC-20.

C.1 The PATHFINDER knowledge base

The basic building block of the PATHFINDER knowledge base is the disease profile or ***frame***. The disease frame consists of ***features*** useful for diagnosis of lymph node diseases. Currently these features include histopathologic findings seen in both low- and high-power magnifications. Each feature is associated with a list of exhaustive and mutually exclusive ***values***. For example, the feature ***pseudofollicularity*** can take on any one of the values ***absent***, ***slight***, ***moderate***, or ***prominent***. These lists of values give the program access to ***severity*** information. In addition, these lists eliminate obvious

interdependencies among the values for a given feature. For example, if pseudofollicularity is ***moderate***, it cannot also be ***absent***.

Evoking strengths and frequencies are associated with each feature-value pair in a disease profile. We are experimenting with different scales for scoring each feature-value pair, and several methods for combining the scores to form a differential diagnosis. A disease-independent import is also assigned to each feature-value but only a two-valued scale is used. This is because, in PATHFINDER, imports are only used to make boolean or yes/no decisions (see below). In addition to import, PATHFINDER utilizes the concept of ***classic*** features for a disease -- within each disease frame, the pathologist marks those feature-value pairs which are considered to be part of the classic pattern of the disease.

The PATHFINDER knowledge base contains information about obvious association between features. This information is of the form: "Don't ask about feature x unless feature y has certain values." For example, it wouldn't make sense to ask about the degree or range of follicularity if there are no follicles in the tissue section. The feature links also serve to identify interdependencies among features. Feature interdependence is a problem because it can lead to inaccuracies in scoring hypotheses.

The prototype knowledge base was constructed by Dr. Nathwani. During the beginning part of 1984, we organized two meetings of the entire team of experts to define the selection of diseases to be included in the system, and the choice of features to be used in the scoring process. After the features are defined (with text, diagrams, and/or slides) we will proceed with the scoring process.

D. Publications Since January 1983

No publications directly related to PATHFINDER. See publications under ONCOCIN for a selection of recent papers by the computer science collaborators.

E. Funding Support

Research Grant submitted to National Institutes of Health, March, 1984.

Grant Title:"Computer-aided Diagnosis of Malignant Lymph Node Diseases"
Principal Investigator: Bharat Nathwani

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

Because our team of experts are in different parts of the country and the computer scientists are not located at the City of Hope, we envision a tremendous use of SUMEX for communication, demonstration of programs, and remote modification of the knowledge base. The proposal mentioned above was developed using the communication facilities of SUMEX.

B. Sharing and Interaction with Other SUMEX-AIM Projects

Our project depends heavily on the techniques developed by the INTERNIST/CADUCEUS project. Although we have not as yet had direct contacts with the group since the start of the PATHFINDER project, we have been able to utilize information and experience with the INTERNIST program gathered over the years through the AIM conferences and on-line interaction. We expect to re-establish these

contacts in the near future. Our experience with the extensive development of the pathology knowledge base utilizing multiple experts should provide for intense and helpful discussions between our two projects.

C. Critique of Resource Management

The SUMEX resource has provided an excellent basis for the development of a pilot project. The availability of a pre-existing facility with appropriate computer languages, communication facilities (especially the TYMNET network), and document preparation facilities allowed us to make good progress in a short period of time. The management has been very useful in assisting with our needs during the start of this project.

III. RESEARCH PLANS

A. Project Goals and Plans

Collection and refinement of knowledge about lymph node pathology

The pilot computer program suggests diagnosis on 45 common diseases of the lymph node (18 benign, 26 primary malignant, and 1 metastatic) based on 77 histologic features. We plan to dramatically increase quantitatively and qualitatively the knowledge base of the system. We will explore the problems of combining knowledge bases created by multiple experts, but based on a common framework.

We also plan to develop techniques for simplifying the acquisition and verification of knowledge from experts, create mapping schemes that will facilitate the understanding of the many classifications of non-Hodgkin's lymphomas. We will also attempt to represent knowledge about special diagnostic entities, such as multiple discordant histologies and atypical proliferations, which do not fit into the classification methods we have utilized.

Representation Research

We hope to enhance the INTERNIST-1 model by structuring features into a useful hierarchy, implementing new methods for scoring hypotheses, creating appropriate explanation capabilities, and formulating and applying high-level heuristics to guide the program.

B. Requirements for Continued SUMEX Use

We are currently dependent on the SUMEX computer for the development of the program. We are in the process of transferring the program over to Portable Standard Lisp, which can then be transferred to the HP9836 workstations available in the Medical Computer Science Group at Stanford. While the switch to workstations will lessen our requirements for computer time for the development of the algorithms, we will continue to need the SUMEX facility for the interaction with each of the research locations specified in our NIH proposal. The HP equipment is currently unable to allow remote access, and thus the program will have to be maintained on the 2060 for use by all non-Stanford users.

C. Requirements for Additional Computing Resources

Most of our computing resources will be met by the 2060 plus the use of the HP9836 workstation. We will need additional file space on the 2060 as we quadruple the size of our knowledge base. We will continue to require access to the 2060 for communication purposes, access to other programs, and for file storage and archiving.

D. Recommendations for Future Community and Resource Development

We encourage the continued exploration by SUMEX of the interconnection of workstations within the mainframe computer setting. We will need to be able to quickly move a program from workstation to workstation, or from workstation back and forth to the mainframe. Software tools that would help the transfer of programs from one type of workstation to another would also be quite useful.

National AIM Project: Computer-Aided Diagnosis of Malignant Lymph Node Diseases

Principal Investigator: Bharat Nathwani, M.D.
Department of Anatomical Pathology
City of Hope National Medical Center
Duarte, California
(213) 359-8111 x 2456 (NATHWANI@SUMEX-AIM)

Lawrence M. Fagan, M.D., Ph.D.
Department of Medicine
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (FAGAN@SUMEX-AIM)

We are building a computer program, called PATHFINDER, to assist in the diagnosis of lymph node pathology. The project is based at the City of Hope National medical center in collaboration with the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 45 common benign and malignant diseases of the lymph node based on 77 histologic features. Our research plans are to develop a full-scale version of the computer program by substantially increasing the quantity and quality of knowledge and to develop techniques for knowledge representation and manipulation appropriate to this application area. The design of the program has been strongly influenced by the INTERNIST/CADUCEUS program developed on the SUMEX resource.

SOFTWARE AVAILABLE ON SUMEX

PATHFINDER-- A version of the PATHFINDER program is available for experimentation on the DEC 2060 computer. This version is a pilot version of the program, and therefore has not been completely tested.

II.A.4.2. RXDX Project

RXDX Project

Robert Lindsay, Ph.D.
Michael Feinberg, M.D., Ph.D.
Manfred Kochen, Ph.D.
University of Michigan
Ann Arbor, Michigan

Jon Heiser, M.D.
Metropolitan State Hospital
Norwalk, California

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

We are developing a prototype expert system that could act as a consultant in the diagnosis and management of depression. Health professionals would interact with the program as they might with a human consultant, describing the patient, receiving advice, and asking the consultant about the rationale for each recommendation. The program will use a knowledge base constructed by encoding the clinical expertise of a skilled psychiatrist in a set of rules. It will use this knowledge base to decide on the most likely diagnosis (endogenous or nonendogenous depression), assess the need for hospitalization, and recommend specific somatic treatments when this is indicated (e.g., tricyclic antidepressants). The treatment recommendation will take into account the patient's diagnosis, age, concurrent illnesses, and concurrent treatments (drug interactions).

B. Medical Relevance and Collaboration

There has been a growing emphasis in American psychiatry on careful diagnosis using clearly defined clinical criteria (Feighner, et al., 1972; Spitzer, et al., 1975, 1980; Feinberg and Carroll, 1982, 1983). These efforts have led to several sets of criteria for the diagnosis of psychiatric disorders. The "St. Louis" criteria (Feighner, et al., 1972) were succeeded by the Research Diagnostic Criteria (RDC), formulated by researchers from St. Louis and New York (Spitzer, et al., 1975). The RDC led directly to the criteria that are now quasi-official in American psychiatry, DSM-III (Spitzer, et al., 1980). All of these criteria lists were based on a combination of clinical opinion and literature review, and use a decision-tree approach to making a diagnosis. These diagnostic systems have been shown to be acceptably reliable, but their validity remains untested. Other groups have used a multivariate statistical approach to diagnosis. Roth and his colleagues (Carney, et al., 1965) published a discriminant index for distinguishing "endogenous" from "neurotic" depressed patients. This work was repeated by Kiloh, et al. (1972) with much the same results, confirming the findings of Carney, et al. (1965).

We have done similar work, deriving two discriminant indices for separating endogenous depressed patients (unipolar or bipolar) from nonendogenous (neurotic) patients. We cross-validated these indices in separate groups of patients, and also validated them against an external standard, the dexamethasone suppression test (Feinberg and Carroll, 1982, 1983). At the same time, we and others have been further

developing this and other biological measures that may differentiate between patients with endogenous and nonendogenous depression. These include neuroendocrine tests such as the dexamethasone suppression test (DST) and quantitative studies of sleep using EEG. Carroll, et al. (1981) have shown that the DST is abnormal in about 67% of patients with endogenous depression (melancholia) and only 5-10% with nonendogenous (neurotic) depression. Kupfer, et al. (1978) and Feinberg, et al. (1982) have similar results with EEG studies of sleep. These biological markers may be useful for routine clinical use, and can certainly be used as external validating criteria to test the performance of different clinical diagnostic methods, including those mentioned above. Furthermore, we have developed biological criteria for "definitely endogenous" depression and "definitely nonendogenous" depression based on DST and sleep EEG. (Carroll, et al., 1980). Our goal is to use these criteria as an external validating criterion for assessing the performance of various new or different diagnostic schemes, in particular an expert system of the sort we are developing.

C. Highlights of Research Progress

This project began in November 1983. We have been examining two other SUMEX-based psychiatry projects, the BLUEBOX project of Mulsant and Servan-Schreiber (1984), and the HEADMED project of Heiser and Brooks (1978, 1980). Mulsant and Servan-Schreiber visited us at Michigan and discussed the rationale and progress of their project. Heiser also visited with us and has agreed to collaborate with our project as a consultant. He is working on psychopharmacology and is attempting to develop and integrate an appropriate knowledge base for our system.

At Michigan, we have encoded most of the Hamilton Rating Scale (Hamilton, 1967) into EMYCIN rules. This is the standard scale (in English) for rating the severity of depression, and many of the items in it will be relevant to our consultant program. We expect to finish this subproject within the next few weeks.

We have begun to collect video recordings of patient interviews. We select patients recently admitted to the University of Michigan Clinical Studies Unit. They are interviewed by Feinberg and the interviews are observed by Lindsay plus a group of psychiatric residents, psychiatrists and psychologists. After the interview, Feinberg is debriefed by Lindsay, and then the others discuss the case. These data will be the initial source of the expert knowledge base for our consultant.

D. List of Relevant Publications

This project has not yet produced any publications. The following list contains the references cited above, including our previous publications relevant to the RxDx project.

1. Carney, M. W. P., Roth, M. and Garside, R. F.:*The diagnosis of depressive syndromes and the prediction of ECT response*, Brit. J. Psychiatry, 111, 659-674, 1965.

2. Carroll, B. J., Feinberg, M., Greden, J. F., Haskett, R. F., James, N. McI., Steiner, M., and Tarika, J. *Diagnosis of endogenous depression: Comparison of clinical, research, and neuroendocrine criteria*, J. Affect Dis., 2, 177-194, 1980.

3. Carroll, B. J., Feinberg, M., Greden, J. F., Tarika, J., Albala, A. A., Haskett, R. F., James, N. McI., Kronfol, Z., Lohr, N., Steiner, M., de Vigne, J-P, and Young, E.:*A specific laboratory test for the diagnosis of melancholia*, Standardization, validation, and clinical utility. Arch. Gen. Psychiatry, 38, 15-22, 1981.

4. Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., and Munoz, R.: *Diagnostic criteria for use in psychiatric research*, Arch. Gen. Psychiatry, 26, 57-63, 1972.

5. Feinberg, M. and Carroll, B. J.: *Separation of subtypes of depression using discriminant analysis: I. Separation of unipolar endogenous depression from non-endogenous depression*, Brit. J. Psychiatry, 140, 384-391, 1982.

6. Feinberg, M. and Carroll, B. J.:*Separation of subtypes of depression using discriminant analysis. II. Separation of bipolar endogenous depression from nonendogenous ("neurotic") depression*, J. Affective Disorders, 5, 129-139, 1983.

7. Feinberg, M. and Carroll, B.J.: *Biological markers for endogenous depression in series and parallel*, Biological Psychiatry 19:3-11, 1984.

8. Feinberg, M. and Carroll, B.J.: *Biological and nonbiological depression*, Presented at Annual Meeting of the Society of Biological Psychiatry, Los Angeles, May, 1984, Abstract #81.

9. Feinberg, M., Gillin, J. C., Carroll, B. J., Greden, J. F., and Zis, A. P.:*EEG studies of sleep in the diagnosis of depression* Biological Psychiatry, 17, 305-316, 1982.

10. Heiser, J. F. and Brooks, R. E.:*Design considerations for a clinical psychopharmacology advisor*, Proc. Second Annual Symp. on Computer Applications in Medical Care. New York: IEEE, 1978, 278-285.

11. Heiser, J. F. and Brooks, R. E.:*Some experience with transferring the MYCIN system to a new domain*, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-2, No. 5, 477-478, 1980.

12. Kiloh, L. G., Andrews, G., and Neilson, M.:*The relationship of the syndromes called endogenous and neurotic depression*, Brit. J. Psychiatry, 121, 183-196, 1972.

13. Kupfer, D. J., Foster, F. G., Coble, P., McPartland, R. J., and Ulrich, R. F.:*The application of EEG sleep for the differential diagnosis of affective disorders*, Am. J. Psychiatry, 135, 69-74, 1978.

14. Mulsant, B. and Servan-Schreiber, D.:*Knowledge engineering: A daily activity on a hospital ward*, Computers in Biomedical Research, 1984.

15. Spitzer, R. L., Endicott, J. and Robins, E.: *Research diagnostic criteria*, (2d ed.) New York State Department of Mental Hygiene, New York Psychiatric Institute, Biometrics Research Division, 1975.

16. Spitzer, R. L.: (Ed.).*Diagnostic and statistical manual of mental disorders*, (3d ed.). Washington, D. C.: American Psychiatric Association, 1980.

17. Van Melle, W.:*The EMYCIN Manual*, Computer Science Department, Stanford University, Report HPP-81-16, 1981.

E. Funding Support

We have submitted an application for support to the Vice-President for Research at the Univ of Michigan, who has funds for "seed money" for faculty research (Total

Direct Costs = $3215). We have prepared a grant application, to be sent to the NIH "Small Grants" Program for the May 1, 1984 deadline (Total Direct Costs = $13,850). These funds should enable us to gather the pilot data we will need as part of a major grant application.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaboration and Program Dissemination via SUMEX

We are collaborating via SUMEX with Dr. Jon Heiser, who worked with Ruven Brooks on HEADMED in the late 1970's. We are sharing a common SUMEX account, and communicating using computer mail. Dr. Heiser will write the section of the expert system dealing with the treatment of depression (and eventually of other psychiatric disorders) while Drs. Feinberg and Lindsay work on the diagnostic parts of the system.

B. Sharing and Collaboration with other SUMEX-AIM Projects

We are also collaborating, although more loosely, with Messrs. Benoit Mulsant and David Servan-Schreiber. They wrote an expert system (BLUEBOX) for the diagnosis and treatment of depression which was a first step in the direction we are going. We have access to BLUEBOX through SUMEX, and have been able to learn from its successes and failures. Ben and David will, we expect, be able to offer us many helpful suggestions on our expert system (RXDX) as they pursue their training in Psychiatry and continue their work in AI in medicine.

C. Critique of Resource Management

We have been using EMYCIN to set up our knowledge base, and have found this program invaluable, since it has saved us many hours of programming in LISP. There are some problems with EMYCIN, many of which center around discrepancies between the the version of EMYCIN described in the manual and the version actually running on SUMEX. We would suggest that EMYCIN be more strongly supported than is now the case, if it and SUMEX are to be even more useful to beginners in AI in Medicine. This may involve added expense, such as would be involved in the purchase of an updated version of EMYCIN, but we would certainly be able to make use of the updated version.

SUMEX itself has been invaluable. We don't have easy access to any other machine of equal computing power which also has a strongly supported LISP available. Specifically, the Dandelion LISP machine at Michigan is not easily accessible, while the LISP compiler available on the Amdahl 5860 here differs from those used at major AI centers such as Stanford and MIT. We have also made good use of the ARPANET connections that SUMEX offers. Feinberg will spend a month of his sabbatical working with Prof. Peter Szolovits at MIT, learning about AI in Medicine. (This is an obvious and necessary step for any physician wanting to begin work in the field.) This visit was arranged using computer mail through SUMEX. Lindsay and Feinberg will be able to continue their collaborative work while the latter is in Cambridge, using the same medium. The alternative would be days lost in the mails and many dollars spent on phone calls. We have also been able to get rapid help with problems that arise with EMYCIN using computer mail, saving days and/or dollars.

III. RESEARCH PLAN

A. Project Goals and Plans

Our immediate objective is to develop an expert system which can differentiate patients with the various subtypes of depressive disorder, and prescribe appropriate

treatment. This system should perform at about the level of a board-certified psychiatrist, i.e. better than an average resident but not as well as a human expert in depression. Eventually, we plan to enlarge the knowledge base so that the expert system can diagnose and prescribe for a wider range of psychiatric patients, particularly those with illnesses which are likely to respond to psychopharmacological agents. We will design the system so that it could be used by non-medical clinicians or by non-psychiatrist MD's as an adjunct to consultation with a human expert.

B. Justification and Requirements for continued SUMEX use

This project is entirely dependent on access to SUMEX. We are using the EMYCIN system on SUMEX. That software is not available to us anywhere else. We also make extensive use of SUMEX as a means of communication and file-sharing with our consultant, Jon Heiser, and with David Mulsant and Benoit Servan-Schreiber. The access to SUMEX resources is essentially our sole means of maintaining contact with the community of researchers working on applications of AI in medicine.

We anticipate that our requirements for computing time and file space will continue to grow as the system evolves.

C. Needs and Plans for Other Computing Resources

As our project evolves and we run into the limitations of EMYCIN and the time-shared SUMEX facility, we anticipate employing different expert systems software. At this time, we are not at a stage to say exactly what that will be, but our project is not sufficiently large that we will be able to mount such a software development project ourselves, so we will depend on development and support elsewhere. Ultimately, when our consultant is made available for field trials and clinical use, it will need to be transported to a personal computer that is large enough support the system yet inexpensive enough to be widely available. A LISP machine is an obvious candidate. While current prices of the necessary hardware are too high, computer prices are continuing to drop. Our design strategy is to avoid limiting ourselves and our aspirations to that which is affordable today; instead we will attempt to project the growth of our project and the price-performance curve of computing such that they meet at some reasonable point in the future.

D. Recommendations for Future Community and Resource Development

Valuable as the present SUMEX facilities are to us, they are in many ways limited and awkward to use. The need for more and more computer cycles and memory continues to grow, of course. However, the major limitation we feel is the difficulty and sometimes the impossibility of making contact with everyone who could be of value to us. We hope that greater emphasis will be put on internetwork gateways. It is important not only to establish more of these, but to develop consistent and convenient standards for electronic mail, electronic file transfers, graphic information transfer, national archives and data bases, and personal filing and retrieval (categorization) systems. The present state of the art is quite limiting, now that the basic concepts of computer networking have become available and have proved their potential.

II.B. Books, Papers, and Abstracts

Publications for the various collaborative projects are summarized in their respective progress reports. They also have been submitted separately on the Scientific Subproject Form IIB. They are not reproduced here to avoid redundancy.

II.C. Resource Summary Table

Detailed resource usage information is summarized starting on page 30. Tabulations of this information also have been submitted separately on the requested Scientific Subproject Form. These are not reproduced here to avoid redundancy.

Appendix A

AIM Management Committee Membership

Following are the current membership lists of the various SUMEX-AIM management committees:

AIM Executive Committee:

FEIGENBAUM, Edward A., Ph.D. (Chairman)
Principal Investigator - SUMEX
Heuristic Programming Project
Department of Computer Science
Margaret Jacks Hall
Stanford University
Stanford, California 94305
(415) 497-4879

LEDERBERG, Joshua, Ph.D.
President
The Rockefeller University
1230 York Avenue
New York, New York 10021
(212) 570-8080, 570-8000

KULIKOWSKI, Casimir, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-2006

LINDBERG, Donald A.B., M.D. (Adv Grp Chrmn)
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

SHORTLIFFE, Edward H., M.D., Ph.D.
Co-Principal Investigator - SUMEX
Division of General Internal Medicine, TC117
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

AIM Advisory Group:

LINDBERG, Donald A.B., 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. (Exec. Secretary)
Biotechnology Resources Program
National Institutes of Health
Building 31, Room 5B43
9000 Rockville Pike
Bethesda, Maryland 20205
(301) 496-5411

FEIGENBAUM, Edward A., Ph.D. (Ex-officio)
Principal Investigator - SUMEX
Heuristic Programming Project
Department of Computer Science
Margaret Jacks Hall
Stanford University
Stanford, California 94305
(415) 497-4879

KULIKOWSKI, Casimir, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-2006

LEDERBERG, Joshua, Ph.D.
President
The Rockefeller University
1230 York Avenue
New York, New York 10021
(212) 570-8080, 570-8000

MINSKY, Marvin, Ph.D.
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139
(617) 253-5864

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 20205
(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

SHORTLIFFE, Edward H., M.D., Ph.D. (Ex-officio)
Co-Principal Investigator - SUMEX
Division of General Internal Medicine, TC117
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

SIMON, Herbert A., Ph.D.
Department of Psychology
Baker Hall, 339
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213
(412) 578-2787, 578-2000

Stanford Community Advisory Committee:

FEIGENBAUM, Edward A., Ph.D. (Chairman)
Heuristic Programming Project
Department of Computer Science
Margaret Jacks Hall
Stanford University
Stanford, California 94305
(415) 497-4879

DJERASSI, Carl, Ph.D.
Department of Chemistry, Stauffer I-106
Stanford University
Stanford, California 94305
(415) 497-2783

MAFFLY, Roy H., M.D.
Division of Nephrology
Veterans Administration Hospital
3801 Miranda Avenue
Palo Alto, California 94304
(415) 858-3971

SHORTLIFFE, Edward H., M.D., Ph.D.
Co-Principal Investigator - SUMEX
Division of General Internal Medicine, TC117
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

Appendix B

Community Growth and Project Abstracts

This appendix contains a graphical display of the development of the SUMEX-AIM community over the years and abstracts of currently active projects. Figure 15 below illustrates the substantial growth in the cumulative number of projects in the Stanford, National AIM, and Rutgers-AIM communities since the resource began operation in 1974 up until this past year. The recent decrease in the total number of projects is due to the closure of several long time SUMEX-AIM projects, namely Dendral, Puff/Vm, Act, and Protein. Activity in the community however remains high, as evidenced by the number of pilot projects (5 Stanford pilots, 2 Aim pilots, and 1 Rutgers pilot) currently active in the SUMEX-AIM community.

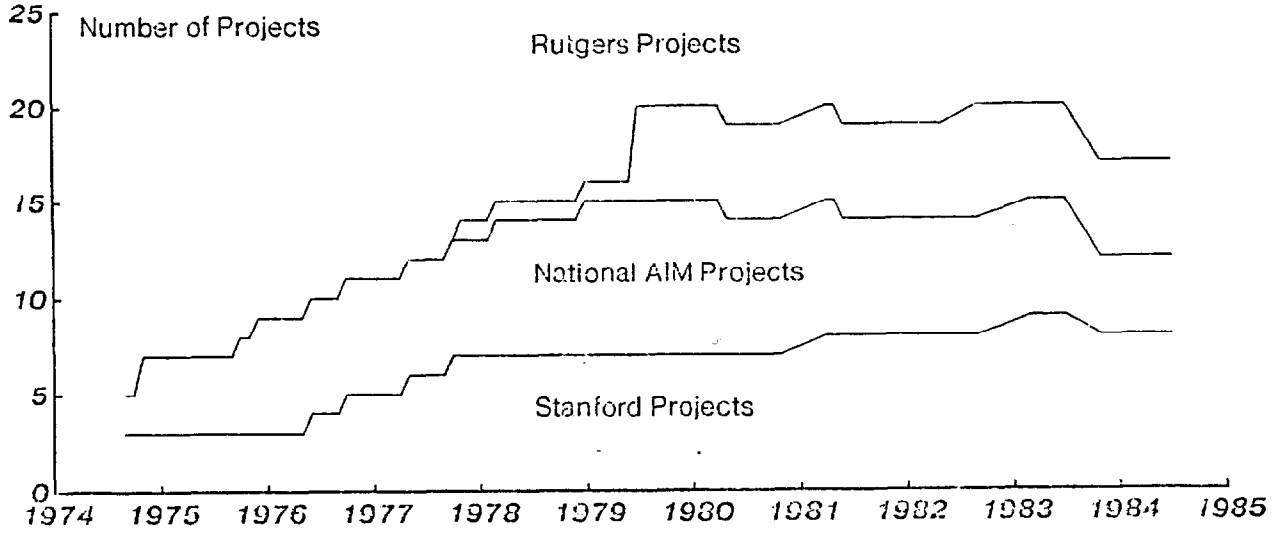

Figure 15: SUMEX-AIM Growth by Community

National AIM Project: CADUCEUS (formerly INTERNIST)

Principal Investigators: Jack D. Myers, M.D. (MYERS@SUMEX-AIM)
Harry E. Pople, Ph.D. (POPLE@SUMEX-AIM)
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
Dr. Pople: (412) 624-3490

The major goal of the CADUCEUS Project is to produce a reliable and adequately complete diagnostic consultative program in the field of internal medicine. Although this program is intended 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. In the design of CADUCEUS and its predecessor INTERNIST I, we have attempted to model the creative, problem-formulation aspect of the clinical reasoning process. The program employs a novel heuristic procedure that composes differential diagnoses, dynamically, on the basis of clinical evidence. During the course of a CADUCEUS or INTERNIST I consultation, it is not uncommon for a number of such conjectured problem foci to be proposed and investigated, with occasional major shifts taking place in the program's conceptualization of the task at hand.

SOFTWARE AVAILABLE ON SUMEX

Versions of INTERNIST are available for experimental use, but the project continues to be oriented primarily towards research and development; hence, a stable production version of the system is not yet available for general use.

REFERENCES

Pople, H.E., Myers, J.D. and Miller, R.A.: *The DIALOG model of diagnostic logic and its use in internal medicine.* Proc. Fourth IJCAI, Tbilisi, USSR, September, 1975.

Pople, H.E.: *The formation of composite hypotheses in diagnostic problem solving: An exercise in synthetic reasoning.* Proc. Fifth IJCAI, Boston, August, 1977.

National AIM Project: SECS -- SIMULATION AND EVALUATION OF CHEMICAL SYNTHESIS

Principal Investigator: W. Todd Wipke, Ph.D.
Department of Chemistry
University of California at Santa Cruz
Santa Cruz, California 95064
(408) 429-2397 (WIPKE@SUMEX-AIM)

The SECS Project aims at developing practical computer programs to assist investigators in designing syntheses of complex organic molecules of biological interest. Key features of this research include the use of computer graphics to allow chemist and computer to work efficiently as a team, the development of knowledge bases of chemical reactions, and the formation of plans to reduce the search for solutions. SECS is being used by the pharmaceutical industry for designing syntheses of drugs.

A spin-off project, XENO, is aimed at predicting the plausible metabolites of foreign compounds for carcinogenicity studies. First, the metabolism is simulated; then the metabolites are evaluated for possible carcinogenicity.

SOFTWARE AVAILABLE ON SUMEX

SECS-- An organic synthesis design program available with a reaction library of over 500 reactions. The program is accessible to users via a teletype or DEC GT40 type graphics terminal.

XENO-- A program for prediction of metabolites of xenobiotic compounds. Although the project is still in the early development stages, this program is available for preliminary exploration and testing.

PRXBLD-- A facility for building approximate 3-dimensional molecular models from their 2-dimensional representations. The program employs an energy minimization approach and is available both stand-alone and as part of SECS.

QED-- A domain-independent inference engine which represents knowledge in first order predicate calculus.

FSECS-- A forward-working synthesis prototype program for finding starting material oriented syntheses.

SST-- A program for searching through a library of possible starting materials to suggest potential starting materials for a given target molecule.

REFERENCES

Wipke, W.T., Rogers, D.: *Rapid Subgraph Search Using Parallelism.* J. Chem. If. Comput. Sci. (Submitted April 24, 1984).

Wipke, W.T.: *An Integrated System for Drug Design*, in COMPUTERS A-Z: A Manufacturer's Guide to Hardware and Software for the Pharmaceutical Industry, Aster Publishing C., Springfield, Oregon. (In press)

Wipke, W.T., and Rogers, D.: *Artificial Intelligence in Organic*

Synthesis. SST: Starting Material Selection Strategies. An Application of Superstructure Search. J. Chem. Inf. Comput. Sci., 24:0000, 1984.

Wipke, W.T., Ouchi, G.I. and Chou, J.T.: *Computer-assisted prediction of metabolism.* IN L. Goldberg (Ed.), STRUCTURE-ACTIVITY CORRELATIONS AS A PREDICTIVE TOOL IN TOXICOLOGY. Hemisphere Publishing Corp., New York, 1983, pp 151-169.

Wipke, W.T., Ouchi, G. and Krishnan, S.: *Simulation and evaluation of chemical synthesis--SECS. An application of artificial intelligence techniques.* Artificial Intelligence 10:999, 1978.

National AIM Project: CLIPR -- HIERARCHICAL MODELS OF HUMAN COGNITION

Principal Investigators: Walter Kintsch, Ph.D. (KINTSCH@SUMEX-AIM)
Peter G. Polson, Ph.D. (POLSON@SUMEX-AIM)
Computer Laboratory for Instruction
in Psychological Research (CLIPR)
Department of Psychology
University of Colorado
Boulder, Colorado 80302
(303) 492-6991

Contact: Dr. Peter G. Polson (Polson@SUMEX-AIM)

The CLIPR Project is concerned with the modeling of complex psychological processes. It is comprised of two research groups. The prose comprehension group has completed a project that carries out the microstructure text analysis described by Miller and Kintsch (1980), yielding predictions of the recall and readability of that text by human subjects. More recently, this group has been interacting with the Heuristic Programming Project at Stanford, using the AGE and UNITS packages to build a more complex model of the knowledge-based processes characteristic of prose comprehension. The planning group is working toward a model of the planning processes used by expert computer software designers. The initial processes involved in learning to use computers and other complex devices.

SOFTWARE AVAILABLE ON SUMEX

A set of programs has been developed to perform the microstructure text analysis described in Kintsch and van Dijk (Psychological Review, 1978) and Miller and Kintsch (1980). The program accepts a propositionalized text as input, and produces indices that can be used to estimate the text's recall and readability. A more complex model based in AGE and UNITS, which emphasizes the knowledge-based aspects of comprehension, is currently under development.

REFERENCES

Jeffries, R., Turner, A.A., Polson, P.G. and Atwood, M.A.: *The Processes Involved in Designing Software.* IN J.R. Anderson (Ed.), COGNITIVE SKILLS AND THEIR ACQUISITION. Hillsdale, NJ, L. Erlbaum Assoc., 1981. (Forthcoming)

Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity.* Int. J. Man-Machine Studies, In Press.

Kintsch, W.: *On modeling comprehension.* Educ. Psychologist, 14:3-14, 1979.

Miller, J.R. and Kintsch, W.: *Readability and recall of short prose passages: A theoretical analysis.* J. Experimental Psychology: Human Learning and Memory, 1980. (In press)

Rutgers AIM Project: RUTGERS RESEARCH RESOURCE-
COMPUTERS IN BIOMEDICINE

Principal Investigators: Saul Amarel, Ph.D.[1982-83], Casimir Kulikowski, Ph.D.
Sholom M. Weiss, Ph.D.[1983-84]
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546 (AMAREL@RUTGERS)
(201) 932-2006 (KULIKOWSKI@RUTGERS)
(201) 932-2379 (WEISS@RUTGERS)

The Rutgers Research Resource provides the research support with artificial intelligence systems, and the computing support with its DEC2060 facility to a large number of biomedical scientists and researchers. There are currently 86 investigators associated with the Resource. Research activities are concentrated in three major areas: expert medical systems, models for planning and knowledge acquisition, and general AI systems development.

One of the most significant achievements in bringing the work of the Resource to bear on clinical research and practice lies in the transfer of technology from our large DEC20 machine to microprocessor compatible representations. The initial breakthrough came with the automatic translation of a serum protein electrophoresis interpretation model so that a version could be incorporated in an instrument - the scanning densitometer (CliniScan) produced by Helena Laboratories. After testing, it was disseminated commercially, marking the first successful transfer of technology from the Resource to general availability in the clinical community. It is now being used in over one hundred clinical locations.

During the current period, we have started a new project with long term implications for the impact of AIM technology: the development of a hand-held microcomputer version of an expert consultation system for front-line health workers. In collaboration with Dr. Chandler Dawson (UCSF), Director of the World Health Organization's Collaborative Centre for the Prevention of Blindness and Trachoma, we have developed a prototype model for consultation on primary eye care. This has been oriented at problems of injury, infection, malnutrition and cataract in situations where an ophthamologist is unavailable. In most developing nations, the incidence of blindness is 10% to 40% higher than in the USA because of these kinds of problems. With the help of a grant from the USAID, we are developing the systems needed for management of eye disease by front-line health workers in developing nations, and outlying parts of the USA.

Another significant technology transfer experiment involves a very large consultation model. The rheumatology knowledge base developed by our collaborators Drs. Lindberg and Sharp at the University of Missouri has been transferred by us to the MC68000 microprocessor based system, and in the past year testing has begun at their site. This represents a major step in bringing the results of artificial intelligence research to the point where clinical researchers who do not have access to large research machines will be able to make use of the results. We are designing a specialized ***rheumatology machine*** which can carry out the same sophisticated reasoning that now needs the Resource DEC20 , but will cost little over $10,000. Because the transfer has been accomplished we can continue to develop large scale models using the full facilities of the Resource DEC20 , but with the confidence that they can then move out into clinical research environments when completed.

REFERENCES

Weiss, S.M. and Kulikowski, C.A. *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

Kulikowski, C.A. contributor to the Knowledge Acquisition chapter edited by B. Buchanan in the book ***Building Expert Systems*** (F. Hayes-Roth, et al., eds) Addison-Wesley, 1983 (in press).

Yao, Y. and Kulikowski, C.A., ***Multiple Strategies of Reasoning for Expert Systems***, Proc. Sixteenth Hawaii International Conference on Systems Sciences, pp. 510-514 , 1983.*

Kulikowski, C.A.***Progress in Expert AI Medical Consultation Systems: 1980 - 1983*** , Proc. MEDINFO '83 , pp. 499-502, Amsterdam, August 1983.*

Kastner, J.K., Weiss, S.M., and Kulikowski, C.A.,***An Efficient Scheme for Time-Dependent Consultation Systems***, Proc. MEDINFO '83, pp.619-622, 1983.*

Kulikowski, C.A.***Expert Medical Consultation Systems***, Journal of Medical Systems, v.7, pp. 229-234, 1983.*

Weiss, S.M., Kulikowski, C.A., and Galen, R.S.,***Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program***, Journal of Clinical Laboratory Automation, v.3, pp. 383-387, 1983.*

Kastner, J.K., Weiss, S.M., and Kulikowski, C.A., ***An Expert System for Front-Line Health Workers in Primary Eye Care***, Proc. Seventeenth Hawaii International Conference on Systems Sciences, pp. 162-166, 1984.*

Kulikowski, C.A. *Knowledge Acquisition and Learning in EXPERT*, Proc. 1983 Workshop on Machine Learning, Univ. of Illinois,Champaign-Urbana 1983.

Indicate by an asterisk (*) that the resource was given credit.

National AIM Project: SOLVER -- PROBLEM SOLVING EXPERTISE

Principal Investigators: Paul E. Johnson, Ph.D.
Center for Research in Human Learning
205 Elliott Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 373-5302 (PJOHNSON@SUMEX-AIM)

William B. Thompson, Ph.D.
Department of Computer Science
136 Lind Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 373-0132 (THOMPSON@SUMEX-AIM)

The Minnesota SOLVER project focuses upon the development of strategies for discovering and representing the knowledge and skill of expert problem solvers. Although in the last 15 years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only the limited amount of expertise that individuals are able to report in a particular constrained language (e.g., production rules). What is still lacking is a theoretical framework capable of reducing dependence upon the expert's intuition or on the near exhaustive testing of possible organizations. Our methodology consists of: (1) extensive use of verbal thinking aloud protocols as a source of information from which to make inferences about underlying cognitive structures and processes; (2) development of computer models as a means of testing the adequacy of inferences derived from protocol studies; (3) testing and refinement of the cognitive models based upon the study of human and model performance in experimental settings. Currently, we are investigating problem-solving expertise in domains of medicine, physics, engineering, management, and law.

SOFTWARE AVAILABLE ON SUMEX

A redesigned version of the Diagnoser simulation model, named Galen, has been implemented on SUMEX.

REFERENCES

Johnson, P.E.: *What kind of expert should a system be?* J. Medicine and Philosophy, 8:77-97, 1983.

Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic reasoning*. Cognitive Science 5:235-283, 1981.

Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Diagnostic reasoning in software fault localization*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

Johnson, P.E., "The Expert Mind: A New Challenge for the Information scientist," in *Beyond Productivity: Information System Development for Organizational Effectiveness*, Th. M. A. Bemelmans (editor), Elsevier Science Publishers B. V. (North-Holland), 1984.

National AIM Project: Computer-Aided Diagnosis of Malignant Lymph Node Diseases (PATHFINDER)

Principal Investigator: Bharat Nathwani, M.D.
Department of Anatomical Pathology
City of Hope National Medical Center
Duarte, California
(213) 359-8111 x 2456 (NATHWANI@SUMEX-AIM)

Lawrence M. Fagan, M.D., Ph.D.
Department of Medicine
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (FAGAN@SUMEX-AIM)

We are building a computer program, called PATHFINDER, to assist in the diagnosis of lymph node pathology. The project is based at the City of Hope National medical center in collaboration with the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 45 common benign and malignant diseases of the lymph node based on 77 histologic features. Our research plans are to develop a full-scale version of the computer program by substantially increasing the quantity and quality of knowledge and to develop techniques for knowledge representation and manipulation appropriate to this application area. The design of the program has been strongly influenced by the INTERNIST/CADUCEUS program developed on the SUMEX resource.

SOFTWARE AVAILABLE ON SUMEX

PATHFINDER-- A version of the PATHFINDER program is available for experimentation on the DEC 2060 computer. This version is a pilot version of the program, and therefore has not been completely tested.

Stanford Project: EXPEX -- EXPERT EXPLANATION

Principal Investigator: Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (SHORTLIFFE@SUMEX-AIM)

EXPEX is a recent Stanford research project that is involved with the development of new representation schemes to facilitate knowledge acquisition and explanation. This includes not only the study of fundamental representational formalisms but also the encoding of various types of knowledge, such as causal information and user models. The research effort deals with medical domains and is being undertaken on SUMEX or on professional workstations linked to the central resource.

Our interest in explanation derives from the insights we gained in developing explanatory capabilities for the MYCIN system. In that system and its descendants, we were able to generate intelligible explanations by taking advantage of a rule-based representation scheme. The limitations of the justifications generated using MYCIN's explanation capabilities have become increasingly obvious, however, and have led to improved characterization of the kinds of explanation capabilities that must be developed if clinical consultation systems are to be accepted by physicians. EXPEX is devoted to the development of new practical and theoretical insights into this problem.

REFERENCES

Teach, R.L. and Shortliffe, E.H.: ***An analysis of physician attitudes regarding computer-based clinical consultation systems***. Comput. Biomed. Res. 14:542-558, 1981.

Wallis, J.W. and Shortliffe, E.H.: ***Explanatory power for medical expert systems: Studies in the representation of causal relationships for clinical consultations***. Meth. Info. Med. 21:127-136, 1982.

Langlotz, C.P. and Shortliffe, E.H.: ***Adapting a consultation system to critique user plans***. International Journal of Man-Machine Studies, 19:479-496 (1983)

Tsuji, S. and Shortliffe, E.H.: ***Graphical access to the knowledge base of a medical consultation system.*** Proceedings of AAMSI Congress 83, pp. 551-555, San Francisco, Ca., May 1983.

Hasling, D. W., Clancey, W. J., and Rennels, G.:
Strategic explanations for a diagnostic consultation system.
International Journal of Man-Machine Studies, Spring 1984 (in press).

Stanford Project: GUIDON/NEOMYCIN --
KNOWLEDGE ENGINEERING
FOR TEACHING MEDICAL DIAGNOSIS

Principal Investigators: William J. Clancey, Ph.D.
701 Welch Road
Department of Computer Science
Stanford University
Palo Alto, California 94304
(415) 497-1997 (CLANCEY@SUMEX-AIM)

Bruce G. Buchanan, Ph.D.
Computer Science Department
Margaret Jacks Hall
Stanford University
Stanford, California 94305
(415) 497-0935 (BUCHANAN@SUMEX-AIM)

SOFTWARE AVAILABLE ON SUMEX

GUIDON--A system developed for intelligent computer-aided instruction. Although it was developed in the context of MYCIN's infectious disease knowledge base, the tutorial rules will operate upon any EMYCIN knowledge base.

NEOMYCIN--A consulation system derived from MYCIN, with the knowledge base greatly extended and reconfigured for use in teaching. In contrast with MYCIN, diagnostic procedures, common sense facts, and disease hierarchies are factored out of the basic finding/disease associations. The diagnostic procedures are abstract (not specific to any problem domain) and model human reasoning, unlike the exhaustive, top-down approach implicit in MYCIN's medical rules. This knowledge base will be used in the GUIDON2 family of instructional programs.

REFERENCES

Clancey, W.J.: *Overview of GUIDON.* In A. Barr and E.A. Feigenbaum (Eds.), THE HANDBOOK OF ARTIFICIAL INTELLIGENCE, Vol. 2. William Kaufmann Assoc., Los Altos, CA, 1982. (Also to appear in J. of Computer-based Instruction)

Clancey, W.J.: ***Methodology for building an intelligent tutoring system.*** In Kintsch, Polson, and Miller, (Eds.), METHODS AND TACTICS IN COGNITIVE SCIENCE. L. Erlbaum Assoc., Hillsdale, NJ. (In press) (Also STAN-CS-81-894, HPP 81-18)

Clancey, W.J.: ***Acquiring, representing, and evaluating a competence model of diagnosis.*** In Chi, Glaser, and Farr (Eds.), THE NATURE OF EXPERTISE. In preparation. HPP-84-2.

Stanford Project: MOLGEN -- AN EXPERIMENT PLANNING SYSTEM FOR MOLECULAR GENETICS

Principal Investigators: Edward A. Feigenbaum, Ph.D.
Department of Computer Science
Stanford University

Charles Yanofsky, Ph.D. (YANOFSKY@SUMEX-AIM)
Department of Biology
Stanford University
Stanford, California 94305
(415) 497-2413

Contact: Dr. Peter FRIEDLAND@SUMEX-AIM
(415) 497-3728

The goal of the MOLGEN Project is to apply the techniques of artificial intelligence to the domain of molecular biology with the aim of providing assistance to the experimental scientist. Previous work has focused on the task of experiment design. Two major approaches to this problem have been explored, one which instantiates abstracted experimental strategies with specific laboratory tools, and one which creates plans in toto, heavily influenced by the role played by interactions between plan steps. As part of the effort to build an experiment design system, a knowledge representation and acquisition package--the UNITS System, has been constructed. A large knowledge base, containing information about nucleic acid structures, laboratory techniques, and experiment-design strategies, has been developed using this tool. Smaller systems, such as programs which analyze primary sequence data for homologies and symmetries, have been built when needed.

New work has begun on scientific theory formation, modification, and testing. This work will be done within the domain of regulatory genetics. We plan to explore fundamental issues in machine learning and discovery, as well as construct systems that will assist the laboratory scientist in accomplishing his intellectual goals.

SOFTWARE AVAILABLE ON SUMEX

SPEX system for experiment design.
UNITS system for knowledge representation and acquisition.
SEQ system for nucleotide sequence analysis.

REFERENCES

Friedland, P.E.: *Knowledge-based experiment design in molecular genetics,* (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-79-771.

Friedland, P.E.: *Knowledge-based experiment design in molecular genetics,* Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 285-287.

Stefik, M.J.: *An examination of a frame-structured representation system,* Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 845-852.

Stefik, M.J.: *Planning with constraints,* (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-80-784, March, 1980.

Stanford Project: ONCOCIN -- KNOWLEDGE ENGINEERING FOR ONCOLOGY CHEMOTHERAPY CONSULTATION

Principal Investigator: Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (SHORTLIFFE@SUMEX-AIM)

Project Directors: Dr. Lawrence M. Fagan and Ms. Miriam B. Bischoff

The ONCOCIN Project is overseen by a collaborative group of physicians and computer scientists who are developing an intelligent system that uses the techniques of knowledge engineering to advise oncologists in the management of patients receiving cancer chemotherapy. The general research foci of the group members include knowledge acquisition, inexact reasoning, explanation, and the representation of time and of expert thinking patterns. Much of the work developed from research in the 1970's on the MYCIN and EMYCIN programs, early efforts that helped define the group's research directions for the coming decade. MYCIN and EMYCIN are still available on SUMEX for demonstration purposes.

The prototype ONCOCIN system is in routine use by oncologists in the Stanford Oncology Clinic. Thus much of the emphasis of this research has been on human engineering so that the physicians will accept the program as a useful adjunct to their patient care activities. ONCOCIN has been well-accepted since its introduction, and plans are underway to transfer the program to professional workstations (rather than the central SUMEX computer) so that it can be implemented and evaluated at sites away from the University.

SOFTWARE AVAILABLE ON SUMEX

MYCIN-- A consultation system designed to assist physicians with the selection of antimicrobial therapy for severe infections. It has achieved expert level performance in formal evaluations of its ability to select therapy for bacteremia and meningitis. Although MYCIN is no longer the subject of an active research program, the system continues to be available on SUMEX for demonstration purposes and as a testing environment for other research projects.

EMYCIN-- The "essential MYCIN" system is a generalization of the MYCIN knowledge representation and control structure. It is designed to facilitate the development of new expert consultation systems for both clinical and non-medical domains.

ONCOCIN-- This system is in routine use but is designed for special high speed terminals and therefore cannot be tested or demonstrated via network connections. Much of the knowledge in the domain of cancer chemotherapy is already well-specified in protocol documents, but expert judgments also need to be understood and modeled.

REFERENCES

Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campbell, A.B., van Melle, W. and Jacobs, C.D.: *ONCOCIN: An expert system for oncology protocol management.* Proc. Seventh IJCAI, pp. 876-881, Vancouver, B.C., August, 1981.

Suwa, M., Scott, A.C. and Shortliffe, E.H.: *An approach to verifying completeness and consistency in a rule-based expert system.* AI Magazine 3(4):16-21, 1982.

Duda, R.O. and Shortliffe, E.H.: *Expert systems research.* Science 220:261-268, 1983.

Langlotz, C.P. and Shortliffe, E.H.: *Adapting a consultation system to critique user plans.* Int. J. Man-Machine Studies 19:479-496, 1983.

Tsuji, S. and Shortliffe, E.H.: *Graphical access to the knowledge base of a medical consultation system.* Proc. AAMSI Congress, pp. 551-555, San Francisco, CA, May 1983.

Bischoff, M.B., Shortliffe, E.H., Scott, A.C., Carlson, R.W. and Jacobs, C.D.: *Integration of a computer-based consultant into the clinical setting.* Proceedings 7th Annual Symposium on Computer Applications in Medical Care, pp. 149-152, Baltimore, Maryland, October 1983.

Mulsant, B. and Servan-Schreiber, D.: *Knowledge engineering: a daily activity on a hospital ward.* Computers and Biomedical Research 17:71-91 (1984).

Stanford Project: RADIX -- DERIVING KNOWLEDGE FROM TIME-ORIENTED CLINICAL DATABASES

Principal Investigators: Robert L. Blum, M.D.
Departments of Medicine
and Computer Science
Stanford University
Stanford, California 94305
(415) 497-9421 (BLUM@SUMEX-AIM)

Gio C.M. Wiederhold, Ph.D.
Department of Computer Science
Stanford University
Stanford, California 94305
(415) 497-0685 (WIEDERHOLD@SUMEX-AIM)

The objective of clinical database (DB) systems is to derive medical knowledge from the stored patient observations. However, the process of reliably deriving causal relationships has proven to be quite difficult because of the complexity of disease states and time relationships, strong sources of bias, and problems of missing and outlying data.

The goal of the RADIX Project is to explore the usefulness of knowledge-based computational techniques in solving this problem of accurate knowledge inference from non-randomized, non-protocol patient records. Central to RADIX is a knowledge base (KB) of medicine and statistics, organized as a taxonomic tree consisting of frames with attached data and procedures. The KB is used to retrieve time-intervals of interest from the DB and to assist with the statistical analysis. Derived knowledge is incorporated automatically into the KB. The American Rheumatism Association DB containing records of 1700 patients is used.

SOFTWARE AVAILABLE ON SUMEX

RADIX--(excluding the knowledge base and clinical database) consists of approximately 400 INTERLISP functions. The following groups of functions may be of interest apart from the RADIX environment:

SPSS Interface Package -- Functions which create SPSS source decks and read SPSS listings from within INTERLISP.

Statistical Tests in INTERLISP -- Translations of the Piezer-Pratt approximations for the T,F, and Chi-square tests into LISP.

Time-Oriented Data Base and Graphics Package -- Autonomous package for maintaining a time-oriented database and displaying labelled time-intervals.

REFERENCES

Monograph, Doctoral Thesis

Blum, R.L.: ***Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project.***
IN D.A.B. Lindberg and P.L. Reichertz (Eds.), LECTURE NOTES IN MEDICAL INFORMATICS, Vol. 19, Springer-Verlag, New York, 1982.

Blum, R.L.: *Discovery and representation of causal relationships from a large time-oriented clinical database: The RX project* (Ph.D. thesis). Computer Science and Biostatistics, Stanford University, 1982.

Journal Articles

Blum, R.L.: *Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical database: The RX Project*. Computers and Biomedical Research 15(2):164-187, April, 1982.

Blum, R.L.: *Displaying clinical data from a time-oriented database.* Computers in Biology and Medicine 11(4):197-210, 1981.

Yu, V., Fagan, L., Wraith S., Clancey, W., Scott, A., Hannigan, J., Blum, R., Buchanan, B. and Cohen, S.: *Antimicrobial selection by computer: A blinded evaluation by infectious diseases experts.* JAMA 242(12):1279-1282, September, 1979.

Conference Proceedings

Blum, R.L. and Wiederhold, G.: *Studying hypotheses on a time-oriented database: An overview of the RX project.* Proc. Sixth SCAMC, IEEE, Washington, D.C., October, 1982.

Blum, R.L.: *Induction of causal relationships from a time-oriented clinical database: An overview of the RX project.* Proc. AAAI Pittsburgh, PA. August, 1982.

Blum, R.L.: *Automated induction of causal relationships from a time-oriented clinical database: The RX Project.* Proc. AMIA Congress, San Francisco, CA, 1982.

Blum, R.L.: *Automating the study of clinical hypotheses on a time-oriented database: The RX project.* Proc. MEDINFO80, Tokyo, October, 1980, pp. 456-460. (Also Stanford Computer Science Report STAN-CS 79-816)

Blum, R.L. and Wiederhold, G.: *Inferring knowledge from clinical data banks utilizing techniques from artificial intelligence.* Proc. Second SCAMC, IEEE, Washington, D.C., November, 1978.

Drafts Submitted for Publication

Blum, R.L.: *Clinical decision making aboard the Starship Enterprise.* Chairman's paper, Session on Artificial Intelligence and Clinical

Decision Making, Congress of the American Association of Medical Systems and Informatics, San Francisco, May, 1983.

Blum, R.L.: *Representation of empirically derived causal relationships.* IJCAI, Karlsruhe, West Germany, August, 1983.

Blum, R.L.: *Machine representation of clinical causal relationships.* MEDINFO 83, Amsterdam, August, 1983.

Blum, R.L.: *Modeling and encoding clinical causal relationships.* SCAMC, Baltimore, MD, October, 1983.

Stanford Pilot Project: THE COMPUTER-AIDED DECISION ANALYSIS (CAMDA) PROJECT

Co-Principal Investigator: Samuel Holtzman
Ronald A. Howard
Department of Engineering-Economic Systems
Stanford University
Stanford, California 94305

Contact: Samuel Holtzman(HOLTZMAN@SUMEX-AIM)
(415) 497-0486

The CAMDA project is a program of research in the area of medical decision making. The main focus of this effort is to combine decision analysis and artificial intelligence to develop systems that support medical decisions.

Nearly two decades of experience in the application of decision analysis to problems in industry and government have shown that the technique constitutes an extremely helpful tool for making difficult choices. The potential benefit of decision analysis is particularly great when choices must be made in the presence of uncertainty and when the stakes involved are high. This situation is common in medical decisions.

Partly as a result of the high cost of an individual decision analysis, and partly due to the inherent complexity of making choices which involve outcomes such as pain and death, medical decision analysis has remained essentially within the realm of the academic community. Therefore, the majority of patients and physicians have been deprived of the benefits of this powerful technique.

Expert system technology make it possible to bring decision analysis to the medical community in general. By providing a sophisticated modeling methodology, expert systems allow the process of decision analysis (within a specific medical context) to be formalized with sufficient accuracy to make much of the analysis amenable to computer automation. The resulting CAMDA systems could provide an attractive alternative to unaided decision making, and to the usually unaffordable option of analyzing medical decisions individually. Furthermore, these systems can help decision makers think more clearly about the difficult issues they face by providing them with a means to experiment with the logical consequences of their assumptions and preferences.

A major focus of our research effort is the development of RACHEL, an intelligent decision systems for infertile couples. The field of infertility was chosen for several reasons, including the prevalence of the condition, the complexity of the values that are usually attached to the possible outcomes in this field, the rapidly growing set of available tests and treatments, and the time-dependent nature of the human reproductive process.

As part of the development of RACHEL, a substantial portion of the current CAMDA effort is aimed at the development of a general computer-based aid for medical decision analysis, which could be used in other medical decision domains.

References:

Holtzman, S.: *A Model of the Decision Analysis Process,* Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1981.

Holtzman, S.: *A Decision Aid for Patients with End-Stage Renal Disease,* Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

(*) Holtzman, S.:*Intelligent Decision Systems,* Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, forthcoming.

References

1. Feigenbaum, E. A. The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering. Proceedings of the 1978 National Computer Conference, AFIPS Press, 1978.

2. Nilsson, N. J.. *Principles of Artificial Intelligence.* Tioga Publishing Company, Palo Alto, California, 1980.

3. Buchanan, B.G., and E.H. Shortliffe. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programamming Project.* Addison-Wesley, New York, 1984.

4. Winston, P. H.. *Artificial Intelligence.* Addison-Wesley Publishing Co., 1977.

5. Nilsson, N. J.. *Information Processing.* Volume 74: *Artificial Intelligence.* North-Holland Pub. Co., 1975.

6. Barr, A., Cohen, P., and Feigenbaum, E. A. (Eds.). *The Handbook of Artificial Intelligence, Volumes I, II, and III.* William Kaufmann, Inc., Los Altos, California, 1981 and 1982.

7. Boden, M.. *Artificial Intelligence and Natural Man.* Basic Books, New York, 1977.

8. McCorduck, P.. *Machines Who Think.* W. H. Freeman and Co., San Francisco, 1979.

9. Coulter, C. L. "Research Instrument Sharing." *Science 201,* 4354 (August 4, 1978).

10. Hayes-Roth, F., Waterman, D., Lenat, D. (Eds). *Building Expert Systems.* Addison-Wesley, 1983.

11. Metcalfe, R. M. and Boggs, D. R. "Ethernet: Distributed Packet Switching for Local Computer Networks." *Comm. ACM 19,* 7 (July 1976).

12. Shoch, J. F. and Hupp, J. A. Performance of an Ethernet Local Network -- A Preliminary Report. Proceedings of the Local Area Communications Network Symposium, Local Area Communications Network Symposium, Boston, May, 1979.

13. Taft, E. A. Implementation of PUP in TENEX. Internal XEROX PARC memorandum, June 1978.

14. Boggs, D. R., Shoch, J. F., Taft, E. A., and Metcalfe, R. M. PUP: An Internetwork Architecture. Report CSL-79-10, XEROX PARC, July, 1979.

15. Digital Equip. Corp., Intel Corp., and Xerox Corp. The Ethernet - Data Link and Physical Layer Specifications. Version 1.0, September 30, 1980.

16. van Melle, W.. *System aids in constructing consultation programs.* UMI Research Press, Ann Arbor, MI, 1980.