SUMEX

STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE
RR-00785

ANNUAL REPORT—YEAR 14

Submitted to

BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM
NATIONAL INSTITUTES OF HEALTH

June 1, 1987

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

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

DIVISION OF RESEARCH RESOURCES
BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM

ANNUAL PROGRESS REPORT
PART I., TITLE PAGE

1. PHS GRANT NUMBER:	5 P41 RR00785-14
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:	08-01-86
5b. TO:	07-31-87
6. PRINCIPAL INVESTIGATOR:	
6a. NAME:	Edward H. Shortliffe, M.D., Ph.D.
6b. TITLE:	Associate Professor of Medicine and Computer Science
6c. SIGNATURE:	Edward H Shortliffe
7. DATE SIGNED:	May 19, 1987
8. TELEPHONE:	415-723-6979

Table of Contents

I. Title Page **1**

II. Description of Program Activities **3**

II.A. Scientific Subprojects 3
II.B. Books, Papers, and Abstracts 3
II.C. Resource Summary Table 3

III. Narrative Description **5**

III.A. Summary of Research Progress 5
III.A.1. Resource Overview 5
III.A.1.1. SUMEX-AIM as a Resource 5
III.A.1.2. The Future of SUMEX-AIM 7
III.A.2. Resource Definitions and Goals 10
III.A.2.1. Knowledge-Based System Research 10
III.A.2.2. Resource Sharing 11
III.A.2.3. Significance and Impact in Biomedicine 11
III.A.2.4. Summary of Current Resource Goals 12
III.A.3. Details of Technical Progress 16
III.A.3.1. Progress Highlights 16
III.A.3.2. Core ONCOCIN Research 19
III.A.3.3. Core AI Research 25
III.A.3.4. Core System Development 32
III.A.3.5. Relevant Core Research Publications 61
III.A.3.6. Resource Equipment 67
III.A.3.7. Training Activities 77
III.A.3.8. Resource Operations and Usage 79
III.B. Highlights 93
III.B.1. The MENTOR Project 94
III.B.2. The GUIDON Project 95
III.B.3. The PROTEAN Project 96
III.B.4. The Medical Information Science Program 97
III.B.5. Remote Virtual Graphics 98
III.C. Administrative Changes 99
III.D. Resource Management and Allocation 100
III.D.1. Overall Management Plan 100
III.D.2. 2060 Cost Center 101
III.E. Dissemination of Resource Information 103
III.F. Suggestions and Comments 105

IV. Description of Scientific Subprojects **107**

IV.A. Stanford Projects 108
IV.A.1. GUIDON/NEOMYCIN Project 109
IV.A.2. MOLGEN Project 116
IV.A.3. ONCOCIN Project 123

IV.A.4. PROTEAN Project 138
IV.A.5. RADIX Project 146
IV.B. National AIM Projects 159
IV.B.1. INTERNIST-I Project 160
IV.B.2. CLIPR - Hierarchical Models of Human Cognition 165
IV.B.3. MENTOR Project 170
IV.B.4. SOLVER Project 174
IV.B.5. ATTENDING Project 181
IV.C. Pilot Stanford Projects 186
IV.C.1. REFEREE Project 187
IV.D. Pilot AIM Projects 193
IV.D.1. PATHFINDER Project 194
IV.D.2. RXDX Project 200
IV.D.3. Dynamic Systems Project 203
IV.D.4. Knowledge Engineering for Radiation Therapy 211
IV.D.5. Pathophysiologic Diagnosis Project 214
Appendix A. AIM Management Committee Membership 217
Appendix B. Scientific Subproject Abstracts 221
References 245

List of Figures

Figure 1: A Sample OPAL Form 21
Figure 2: TALK Session Example and the Software Layers Involved in TALK 37
Figure 3: File Server Throughput Benchmarks 40
Figure 4: SUMEX-AIM DEC 2060 Configuration 72
Figure 5: SUMEX-AIM Sun File Server Configuration 73
Figure 6: SUMEX-AIM Xerox File Server Configuration 74
Figure 7: SUMEX-AIM VAX File Server Configuration 75
Figure 8: SUMEX-AIM EtherNet Configuration 76
Figure 9: Total CPU Time Consumed by Month 81
Figure 10: Monthly CPU Usage by Community 83
Figure 11: Monthly Terminal Connect Time by Community 84
Figure 12: Cumulative CPU Usage Histogram by Project and Community 86
Figure 13: Public Data Network Terminal Connect Time 90
Figure 14: INTERNET Terminal Connect Time 91
Figure 15: System Downtime -- Hours per Month 92
Figure 16: Overall System Reliability Summary 92
Figure 17: 2060 Cost Center Performance 102

II. Description of Program Activities

This section corresponds to the predefined forms required by the Division of Research Resources to provide information about our resource activities for their computerized retrieval system. These forms have been submitted separately and are not reproduced here to avoid redundancy with the more extensive narrative information about our resource and progress provided in this report.

II.A. Scientific Subprojects

Our core research and development activities are described starting on page 16, our training activities are summarized starting on page 77, and the progress of our collaborating projects is detailed starting on page 107.

II.B. Books, Papers, and Abstracts

The list of recent publications for our core research and development work starts on page 61 and those for the collaborating projects are in the individual reports starting on page 107.

II.C. Resource Summary Table

The details of resource usage, including a breakdown by the various subprojects, is given in the tables starting on page 79.

III. Narrative Description

III.A. Summary of Research Progress

III.A.1. Resource Overview

This is an annual report for year 14 of the SUMEX-AIM resource (grant RR-00785), the first year of a 3-year renewal period to support further research on applications of artificial intelligence in biomedicine. For both technical and administrative reasons, we merged into the June 1985 SUMEX renewal application the continuation of work on the development and dissemination of medical consultation systems (ONCOCIN) that had been supported as resource-related research under grant RR-01631. Progress on core ONCOCIN research is therefore now reported here as well.

These combined efforts represent an ambitious research program to:

- Continue our long-range core research efforts on knowledge-based systems aimed at developing new concepts and methodologies needed for biomedical applications.
- Substantially extend ONCOCIN research on developing and disseminating clinical decision support systems.
- Develop the core system technology to move the national SUMEX-AIM community from a dependence on the central SUMEX DEC 2060 to a fully distributed, workstation-based computing environment.
- Introduce these systems technologies into the SUMEX-AIM community with appropriate communications and managerial assistance to responsibly phase out the central resource and DEC 2060 mainframe in a manner that will support community efforts to become self-sustaining and to continue scientific interactions through fully distributed means.
- Maintain our aggressive efforts at training and dissemination to help exploit the research potential of this field.

III.A.1.1. SUMEX-AIM as a Resource

SUMEX and the AIM Community

In the fourteen years since the SUMEX-AIM resource was established in late 1973, computing technology and biomedical artificial intelligence research have undergone a remarkable evolution and SUMEX has both influenced and responded to these changing technologies. It is widely recognized that our resource has fostered highly influential work in biomedical AI -- work from which much of the expert systems field emerged -- and that it has simultaneously helped define the technological base of applied AI research.

The focus of the SUMEX-AIM resource continues to emphasize research on artificial intelligence techniques that guide the design of computer programs that can help with the acquisition, representation, management, and utilization of the many forms of medical knowledge in diverse biomedical research and clinical care settings -- ranging from biomolecular structure determination and analysis, to molecular biology, to clinical decision support, to medical education. Nevertheless, we have long recognized that the ultimate impact of this work in biomedicine will be realized through its assimilation with the full range of methodologies of medical informatics, such as data bases, biostatistics, human-computer interfaces, complex instrument control, and modeling. From the start, SUMEX-AIM work has been grounded in real-world applications, like systems for the interpretation of mass spectral information about biomolecular structures, chemical synthesis, interpretation of x-ray diffraction data on crystals, cognitive modeling, infectious disease diagnosis and therapy, DNA sequence analysis, experiment planning and interpretation in molecular biology, and medical instruction. Our current work extends this emphasis in application domains such as oncology protocol management, clinical decision support, protein structure analysis, and data base information retrieval and analysis. All of these research efforts have demanded close collaborations with diverse parts of the biomedical research community and the integration of many computational methods from those domains with knowledge-based approaches. Even though in the beginning the "AI-in-medicine" community was quite small, it is perforce no longer limited and easily-defined, but rather is spreading and is inextricably linked with the many biomedical applications communities we have collaborated with over the years. Driven both by the on-going diffusion of AI and by the development of personal computer workstations that signal the practical decentralization of computing resources, we must develop new resource communication and distributed computing technologies that will continue to facilitate wider intra- and inter-community communication, collaboration, and sharing of biomedical information.

The SUMEX Project has demonstrated that it is possible to operate a computing research resource with a national charter and that the services providable over networks were those that facilitate the growth of AI-in-Medicine. SUMEX now has a reputation as a model national resource, pulling together the best available interactive computing technology, software, and computer communications in the service of a national scientific community. Planning groups for national facilities in cognitive science, computer science, and biomathematical modeling have discussed and studied the SUMEX model and new resources, like the recently instituted BIONET resource for molecular biologists, are closely patterned after the SUMEX example.

The projects SUMEX supports have generally required substantial computing resources with excellent interaction. Even today though, with the growing, but by no means ubiquitous availability of workstations, this computing power is still hard to obtain in all but a few universities. SUMEX is, in a sense, a "great equalizer". A scientist gains access by virtue of the quality of his/her research ideas, not by the accident of where s/he happens to be situated. In other words, the resource follows the ethic of the scientific journal.

SUMEX has demonstrated that a computer resource is a useful "linking mechanism" for bringing together and holding together teams of experts from different disciplines who share a common problem focus. AI concepts and software are among the most complex products of computer science. Historically it has not been easy for scientists in other fields to gain access to and mastery of them. Yet the collaborative outreach and dissemination efforts of SUMEX have been able to bridge the gap in numerous cases. Over 36 biomedical AI application projects have developed in our national community and have been supported by SUMEX computing resources over the years. And 9 of these have matured to the point of now continuing their research on facilities outside of SUMEX. For example, the BIONET resource (named GENET while at SUMEX) is being operated by IntelliGenetics; the CADUCEUS project splits their research work between their own IBM PC workstations, a VAX computer, and the SUMEX resource; and the Chemical Synthesis project now operates entirely on a VAX at U.C. Santa Cruz.

The integration of AI ideas with other parts of medical informatics and their dissemination into biomedicine is happening largely because of the development in the

1970's and early 1980's of methods and tools for the application of AI concepts to difficult professional-level problem solving. Their impact was heightened because of the demonstration in various areas of medicine and other life sciences that these methods and tools really work. Here SUMEX has played a key role, so much so that it is regarded as "the home of applied AI."

SUMEX has been the nursery, as well as the home, of such well-known AI systems as DENDRAL (chemical structure elucidation), MYCIN (infectious disease diagnosis and therapy), INTERNIST (differential diagnosis), ACT (human memory organization), ONCOCIN (cancer chemotherapy protocol advice), SECS (chemical synthesis), EMYCIN (rule-based expert system tool), and AGE (blackboard-based expert system tool). In the past four years, our community has published a dozen books that give a scholarly perspective on the scientific experiments we have been performing. These volumes, and other work done at SUMEX, have played a seminal role in structuring modern AI paradigms and methodology.

III.A.1.2. The Future of SUMEX-AIM

Given this background, what is the future need and course for SUMEX as a resource -- especially in view of the on-going revolution in computer technology and costs and the emergence of powerful single-user workstations and local area networking? The answers remain clear.

Basic Research in AI in Biomedicine

At the deepest research level, despite our considerable success in working on medical and biological applications, the problems we can attack are still sharply limited. Our current ideas fall short in many ways against today's important health care and biomedical research problems brought on by the explosion in medical knowledge and for which AI should be of assistance. Just as the research work of the 70's and 80's in the SUMEX-AIM community fuels the current practical and commercial applications, our work of the late 80's will be the basis for the next decade's systems.

The report of the panel on medical informatics [12], convened late in 1985 by the National Library of Medicine to review and recommend twenty-year goals for the NLM, listed among its highest priority recommendations the need to greatly expand and aggressively pursue an interdisciplinary research program to develop computational methods for acquiring, representing, managing, and using biomedical knowledge of all sorts for health care and biomedical research. These are precisely the problems which the SUMEX-AIM community has been working on so successfully and which will require work well beyond the five year funding period we have requested. It is essential that this line of research in the SUMEX-AIM community, represented by our core AI research, the ONCOCIN research, and our collaborative research groups, be continued.

The Changing Role of the Central Resource

At the resource level, there are changing, but still intense, needs for computing resources for the active AIM research community to continue its work over the next five years. The workstations to which we directed our attention in 1980 have now demonstrated their practicality as research tools and, increasingly, as potential mechanisms for disseminating AI systems as cost-effective decision aids in clinical settings.such as private offices. Over the next half decade we expect the era of highly centralized general machines for AI research will come to an end, and be replaced gradually by networks of distributed but heterogeneous single-user machines sharing common information resources and communication paths among members of the biomedical research community.

Many of our community groups are still dependent on the SUMEX-AIM resources. For those that have been able to take advantage of newly developed local computing facilities, SUMEX-AIM provides a central cross-roads for communications and the sharing of programs and knowledge. In its core research and development role, SUMEX-AIM has its sights set on the hardware and software systems of the next decade. We expect major changes in the distributed computing environments that are just now emerging in order to make effective use of their power and to adapt them to the development and dissemination of biomedical AI systems for professional user communities. In its training role, SUMEX is a crucial resource for the education of badly needed new researchers and professionals to continue the development of the biomedical AI field. The "critical mass" of the existing physical SUMEX resource, its development staff, and its intellectual ties with the Stanford Knowledge Systems Laboratory, make this an ideal setting to integrate, experiment with, and export these methodologies for the rest of the AIM community.

At the beginning, the SUMEX community was small and idea-limited, and the central SUMEX computer facility was an ideal vehicle for the research. Now the community is large, and the momentum of the science is such that its progress is limited by computing power and research manpower. The size and scientific maturity of the SUMEX community has fully consumed the computing resource in every critical dimension -- CPU power, main memory size, address space, and file space -- and has overflowed to decentralized machines of many types. Much of our work has already been focussed on developing and experimenting with workstation environments for biomedical AI applications. We are fully committed to continuing this line of research for the future hardware thrust of the resource. We will continue our *experimental* approach to these systems, rejecting articles of faith for real experience. We must learn to build and exploit distributed networks of these machines and to build and manage graceful software for these systems. Since decentralization is central to our future, we must learn its technical characteristics.

The resource development directions we have sketched have received substantial external impetus as well [12, 2, 7]. For example, another of the key recommendations of the NLM medical informatics planning panel [12] was that high-speed network communication links be established throughout the biomedical research community so that knowledge and information can be shared across diverse research groups and that the required interdisciplinary collaborations can take place. A principal goal from the start of SUMEX-AIM has been to experiment with these electronic links, but SUMEX is only a start toward this broad goal. Nevertheless, it continues to be an important pathfinder to develop the technology and community interaction tools needed to expand community system and communication resources.

Highlights of Long-term Goals

- Maintain the synergistic relationship between SUMEX core system development, core AI research, our experimental efforts at disseminating clinical decision-making aids, and new applications efforts.

- Continue to serve the national AIM research community, less and less as a source of raw computing cycles and more and more as a transfer point for new technologies important for community research and communication. We will also continue our coordinating role within the community through electronic media and periodic AIM workshops.

- Maintain our connections to ARPANET, TELENET, and our local Ethernet and assist other community members to establish similar links by example, by integrating and providing enabling software, and by offering advice and support within our resources.

- Focus new computing resource developments on more effective exploitation of distributed workstations through better communication and cooperative computing tools, using transparent digital networking schemes.

- Enhance the computing environments of workstations so that minimal dependency on central, general-purpose computing hosts remains and these mainframe time-sharing systems can be phased out eventually. Remaining central resources will include servers for communications, community information resources, and special computing architectures (e.g., shared- or distributed-memory symbolic multiprocessors) justified by cost-effectiveness and unique functionality.

- Incrementally phase-in, disseminate, and evaluate those aspects of the local distributed computing resource that are necessary for continuing national AIM community support within this distributed paradigm. This will ultimately point the way towards the distributed computing resource model that we believe will interlink this community well into the next decade.

- Gradually and responsibly phase out the existing DEC 2060 machine as effective distributed computing alternatives become widely available. We expect this to be possible sometime during the fourth through fifth years of the continuation resource.

- Continue the central staff and management structure, essentially unchanged in size and function during the five-year transition period, except for the merging of the core part of the ONCOCIN research with the SUMEX resource.

III.A.2. Resource Definitions and Goals

SUMEX-AIM is a national computer resource with a multiple mission: a) promoting experimental applications of computer science research in artificial intelligence (AI) to biological and medical problems, b) studying methodologies for the dissemination of biomedical AI systems into target user communities, c) supporting the basic AI research that underlies applications, and d) facilitating network-based computer resource sharing, collaboration, and communication among a national scientific community of health research projects. The SUMEX-AIM resource 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.

III.A.2.1. Knowledge-Based System Research

The SUMEX Project has given strong impetus to the development of knowledge-based system research in biomedicine. Knowledge-based system research is that part of computer science that investigates symbolic reasoning processes, and the representation of symbolic knowledge for use in inference[1]. A knowledge-based or expert system is a 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.

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 and of good guessing) that characterize expert-level decision-making in the field. Our work views *heuristic* knowledge to be of equal importance with *factual* knowledge, indeed to be the essence of what we call *expertise*. The performance level of an expert system is primarily a function of the size and quality of the knowledge base that it possesses.

Projects in the SUMEX-AIM community are concerned in some way with the application of AI to biomedical research. Brief abstracts of the various projects currently using the SUMEX resource can be found in Appendix B and more detailed progress summaries in Section IV. The most tangible objective of this approach is the development of computer programs that will be more general and effective consultative tools for the clinician and medical scientist. All of these research efforts have demanded close collaborations with diverse parts of the biomedical research community and the integration of many computational methods from those domains with knowledge-based approaches. We have long recognized that the ultimate impact of this work in biomedicine will be realized through its assimilation with the full range of methodologies of medical informatics, including, for example, data base research, biostatistics, decision support, complex instrument control, and modeling.

There have already been promising results in many application areas, even though state-of-the-art programs are far more narrowly specialized and inflexible than the corresponding aspects of human intelligence they emulate. Needless to say, much is yet

[1]Many introductory and survey texts have been written by now on AI and knowledge-based or expert systems. See for example [1, 11, 13, 5, 23, 4, 18].

to be learned in the process of fashioning a coherent scientific discipline out of the experimental programs, mathematical procedures, and emerging theoretical structure comprising knowledge-based system research.

III.A.2.2. Resource Sharing

An equally important function of the SUMEX-AIM resource is an exploration of the use of computer communications as a means for interactions and sharing between geographically remote research groups engaged in biomedical computer science research and for the dissemination of AI technology. This facet of scientific interaction is becoming increasingly important with the explosion of complex information sources and the regional specialization of groups and facilities that might be shared by remote researchers [10, 3]. And, as projected, we are seeing a growing decentralization of computing resources with the emerging technology in microelectronics and a correspondingly greater role for digital communications to facilitate scientific exchange.

Our community building effort is based upon the developing state of distributed computing and communications technology. While far from perfected, these capabilities offer powerful tools for collaborative linkages, both within a given research project and among them. A number of the active projects on SUMEX are based upon the collaboration of computer and medical scientists at geographically separate institutions, separate both from each other and from the computer resource (see for example, the MENTOR and PathFinder projects).

In the early 1970's, the initial model for SUMEX-AIM as a centralized resource was based on the high cost of powerful computing facilities and the infeasibility of being able to duplicate them readily. This central role has already evolved significantly and continues to change with the introduction of more compact and inexpensive computing technology now available at many more research sites. At the same time, the number of active groups working on biomedical AI problems has grown and the established ones have increased in size. This has led to a growth in the demand for computing resources far beyond what SUMEX-AIM could reasonably and effectively provide on a national scale. We have therefore turned our core systems research to actively supporting the development of distributed computing and communications resources to facilitate collaborative project research and continued inter-group communications. Thus, as more remotely available resources have become established, the balance of the use of the SUMEX-AIM resource has shifted toward supporting start-up pilot projects and the growing AI research community at Stanford.

III.A.2.3. Significance and Impact in Biomedicine

Artificial intelligence is the computer science of representations of symbolic knowledge and its use in symbolic inference and problem-solving processes. For computer applications in medicine and biology, this research path is crucial. Medicine and biology are not presently mathematically-based sciences; unlike physics and engineering, they are seldom capable of exploiting the mathematical characteristics of computation. They are essentially inferential, not calculational, sciences. If the computer revolution is to affect biomedical scientists, computers will be used as inferential aids.

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.

Perhaps the larger impact on medicine and biology will be the exposure and refinement of the hitherto largely private heuristic knowledge of the experts of the various fields studied. The ethic of science that calls for the public exposure and criticism of knowledge has traditionally been flawed for want of a methodology to evoke and give form to the heuristic knowledge of scientists. AI methodology is beginning to fill that need. Heuristic knowledge can be elicited, studied, critiqued by peers, and taught to students.

The importance of AI research and its applications is increasing in general, without regard for the specific areas of biomedical interest. AI is one of the principal fronts along which university computer science groups are expanding. The pressure from student career-line choices is great: to cite an admittedly special case, approximately 80% of the students applying to Stanford's computer science Ph.D. program cite AI as a possible field of specialization (up from 30% a few years ago). Federal and industrial support for AI research is vigorous and growing, although support specifically for biomedical applications continues to be limited. All of the major computer manufacturers (e.g., IBM, DEC, TI, UNISYS, HP, and others) are using and marketing AI technology aggressively and many software companies are putting more and more products on the market. Many other parts of industry are also actively pursuing AI applications in their own contexts, including defense and aerospace companies, manufacturing companies, financial companies, and others.

Despite the limited research funding available, there is also an explosion of interest in medical AI. The American Association for Artificial Intelligence (AAAI), the principal scientific membership organization for the AI field, has 7000 members, over 1000 of whom are members of the medical special interest group known as the AAAI-M. Speakers on medical AI are prominently featured at professional medical meetings, such as the American College of Pathology and American College of Physicians meetings; a decade ago, the words *artificial intelligence* were never heard at such conferences. And at medical computing meetings, such as the annual Symposium on Computer Applications in Medical Care (SCAMC) and the international MEDINFO conferences, the growing interest in AI and the rapid increase in papers on AI and expert systems are further testimony to the impact that the field is having.

AI is beginning to have a similar effect on medical education. 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. They all cite evolving computing technology and (SUMEX-AIM) AI research as key motivators. At Stanford, we have vigorous special programs for student training and research in AI -- a new graduate program in Medical Information Sciences and the two-year Masters Degree in AI program. All of these have many more applicants than available slots. Demand for their graduates, in both academic and industrial settings, is so high that students typically begin to receive solicitations one or two years before completing their degrees.

III.A.2.4. Summary of Current Resource Goals

The following outlines the specific objectives of the SUMEX-AIM resource during the current three-year award period begun in August 1986. It provides an overall research plan for the resource and provides the backdrop against which specific progress is reported. Note that these objectives cover only the resource nucleus; objectives for individual collaborating projects are discussed in their respective reports in Section IV. Specific aims are broken into five categories: 1) Technological Research and Development, 2) Collaborative Research, 3) Service and Resource Operations, 4) Training and Education, and 5) Dissemination.

1) Technological Research and Development

SUMEX funding and computational support for core research is complementary to similar funding from other agencies (including DARPA, NASA, NSF, NLM, private foundations, and industry) 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 underpinnings for increasingly effective consultative programs in medicine and for more practical adaptations of this work within emerging microelectronic technologies. Specific aims include:

- Basic research on AI techniques applicable to biomedical problems. Over the next term we will emphasize work on blackboard problem-solving frameworks and architectures, knowledge acquisition or learning, constraint satisfaction, and qualitative simulation.

- Investigate methodologies for disseminating application systems such as clinical decision-making advisors into user groups. This will include generalized systems for acquiring, representing and reasoning about complex treatment protocols such as are used in cancer chemotherapy and which might be used for clinical trials.

- Support community efforts to organize and generalize AI tools and architectures that have been developed in the context of individual application projects. This will include retrospective evaluations of systems like the AGE blackboard experiment and work on new systems such as BB1, MRS, SOAR, EONCOCIN, EOPAL, Meta-ONYX, and architectures for concurrent symbolic computing. The objective is to evolve a body of software tools that can be used to more efficaciously build future knowledge-based systems and explore other biomedical AI applications.

- Develop more effective workstation systems to serve as the basis for research, biomedical application development, and dissemination. We seek to coordinate basic research, application work, and system development so that the AI software we develop for the next 5-10 years will be appropriate to the hardware and system software environments we expect to be practical by then. Our purchases of new hardware will be limited to experimentation with state-of-the-art workstations as they become available for our system developments.

2) Collaborative Research

- Encourage the exploration of 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 consider exceptional applications justified by some other unique feature of SUMEX-AIM essential for important biomedical research. We will continue to exploit community expertise and sharing in software development.

- Minimize administrative barriers to the community-oriented goals of

SUMEX-AIM and direct our resources toward purely scientific goals. We will retain the current user funding arrangements for projects working on SUMEX facilities. User projects will fund their own manpower and local needs; actively contribute their special expertise to the SUMEX-AIM community; and receive an allocation of computing resources under the control of the AIM management committees. We will begin charging "fees for service" to Stanford users as DRR support for the DEC 2060 is phased out. Fees to national users will be delayed as long as financially possible.

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

3) Service and Resource Operations

SUMEX-AIM does not have the computing or manpower capacity to provide routine service to the large community of mature projects that has developed over the years. Rather, their computing needs are better met by the appropriate development of their own computing resources when justified. Thus, SUMEX-AIM has the primary focus of assisting new start-up or pilot projects in biomedical AI applications in addition to its core research in the setting of a sizable number of collaborative projects. We do offer continuing support for projects through the lengthy process of obtaining funding to establish their own computing base.

4) Training and Education

- Provide documentation and assistance to interface users to resource facilities and systems.
- Exploit particular areas of expertise within the community for assisting in the development of pilot efforts in new application areas.
- Accept visitors in Stanford research groups within limits of manpower, space, and computing resources.
- Support the Medical Information Science and MS/AI student programs at Stanford to increase the number of research personnel available to work on biomedical AI applications.
- Support workshop activities including collaboration with other community groups on the AIM community workshop and with individual projects for more specialized workshops covering specific research, application, or system dissemination topics.

5) Dissemination

While collaborating projects are responsible for the development and dissemination of their own AI systems and results, the SUMEX resource will work to provide community-wide support for dissemination efforts in areas such as:

- Encourage, contribute to, and support the on-going export of software systems and tools within the AIM community and for commercial development.

- Assist in the production of video tapes and films depicting aspects of AIM community research.
- Promote the publication of books, review papers, and basic research articles on all aspects of SUMEX-AIM research.

III.A.3. Details of Technical Progress

This section gives an overview of progress for the nucleus of the SUMEX-AIM resource. A more detailed discussion of our progress in specific areas and related plans for further work are presented in Section III.A.3.2. Objectives and progress for individual collaborating projects are discussed in their respective reports in Section IV. These collaborative projects collectively provide much of the scientific basis for SUMEX as a resource and our role in assisting them has been a continuation of that evolved in the past. Collaborating projects are autonomous in their management and provide their own manpower and expertise for the development and dissemination of their AI programs.

III.A.3.1. Progress Highlights

In this section we summarize highlights of SUMEX-AIM resource activities over the past year (May 1986 - April 1987), focusing on the resource nucleus.

- We have made significant progress in the core ONCOCIN research work to generalize the tools for clinical trial management from the initial cancer chemotherapy management application. We began examining the structures of protocols across several medical subspecialties other than cancer chemotherapy, concentrating this year on insulin diabetes treatment. Graphical tools are under development to facilitate protocol definition and knowledge base entry and we worked on model-based reasoning to infer protocol therapeutic actions not explicitly encoded in the decision plan. We have also continued to examine the issues of disseminating the ONCOCIN system into actual clinical settings.

- We made significant progress in core AI research, primarily in the areas of knowledge representation, blackboard frameworks, parallel symbolic computing architectures, and machine learning. Work has advanced on the representation of explicit strategic knowledge for problem-solving and blackboard control knowledge, including cost/benefit trade-offs of increasingly complex control reasoning. The parallel architectures work has developed a flexible, instrumented simulator of distributed-memory, multiprocessor architectures and two alternative parallel blackboard frameworks for expressing application problems. These have been applied to several signal understanding problems with promising nearly linear problem-solving speedup. The machine learning work has concentrated on explanation-based generalization and chunking work in the SOAR framework, inductive rule learning, and tools for debugging knowledge structures. Work has also continued on reasoning with uncertainty to find ways of combining formal and informal approximate reasoning methods. We also continued work on extending and refining the BB1 blackboard system.

- We have made excellent progress on the core system development work targeted at supporting the distributed AIM community. We have continued implementation of uniform network protocol standards for remote workstation access, redirected our virtual graphics work to take advantage of the X window protocol being adopted by many workstation vendors, and implemented prototype communication tools that integrate text and graphics between linked machines. We have concentrated on the NFS protocol for distributed file access and have got experimental versions of this and the underlying remote procedure call facilities working or underway for all of our workstations. An additional service is being implemented to allow remote database queries through remote procedure calls to a standard relational database. We have a prototype distributed electronic mail system working on Xerox D-machines and will be extending and porting this to other environments shortly. We have also made important progress in extending the general computing environments for text processing, file management, printing, communications, and other services on specific workstation environments, including the support of 6 different operating system environments.

- We have continued the dissemination of SUMEX-AIM technology through various media. We have reorganized the distribution system for our AI software tools (EMYCIN, AGE, MRS, SACON, and BB1) to academic, industrial, and federal research laboratories, in order to make it more efficient and require less research staff time. We have also continued to distribute the video tapes of some of our research projects including ONCOCIN, and an overview tape of Knowledge Systems Laboratory work to outside groups. Our group has continued to publish actively on the results of our research, including more than 45 research papers per year in the AI literature and a dozen books in the past 5 years on various aspects of SUMEX-AIM AI research.

- The Medical Information Sciences program, begun at Stanford in 1983 under Professor Shortliffe as Director, has continued its strong development over the past year. The specialized curriculum offered by the MIS program focuses on the development of a new generation of researchers able to support the development of improved computer-based solutions to biomedical needs. The feasibility of this program resulted in large part from the prior work and research computing environment provided by the SUMEX-AIM resource. It has recently received enthusiastic endorsement from the Stanford Faculty Senate for an additional five years, has been awarded renewed post-doctoral training support from the National Library of Medicine with high praise for the training and contributions of the SUMEX-AIM environment from the reviewing study section, and has received additional industrial and foundation grants for student support. This past year, MIS students have published many papers, including several that have won conference awards.

- While the SUMEX-AIM computing resource hardware has been largely unchanged this past year, we continue to evaluate new workstation technologies of advantage to the AIM community. We continue to operate the DEC 2060 mainframe and file servers for the community. Because of the broad mix of research in the SUMEX-AIM community, no single computer vendor can meet our needs so we have undertaken long-term support of a heterogeneous computing environment, incorporating many types of machines linked through multiprotocol Ethernet facilities.

- We have continued to recruit new user projects and collaborators to explore further biomedical areas for applying AI. A number of these projects are built around the communications network facilities we have assembled, bringing together medical and computer science collaborators from remote institutions and making their research programs available to still other remote users. At the same time we have encouraged older mature projects to build their own computing environments thereby freeing up SUMEX resources for newer projects.

- In June 1986, we moved the SUMEX and Medical Computer Science offices into the newly constructed Stanford Medical School Office Building, funded by the university. This space provides us with almost twice the area we previously occupied and it is laid out so as to promote better interactions between our groups and among our students and research staff.

- SUMEX user projects have made good progress in developing and disseminating effective consultative computer programs for biomedical research. These systems provide expertise in areas like cancer chemotherapy protocol management, clinical diagnosis and decision-making, and molecular biology. We have worked hard to meet their needs and are grateful for their expressed appreciation (see Section IV).

III.A.3.2. Core ONCOCIN Research

ONCOCIN is a data-management and therapy-advising program for complex cancer chemotherapy experiments. The development of the system began in 1979, following the successful generalization of MYCIN into the EMYCIN expert system shell. The ONCOCIN project has evolved over the last eight years. The original version of ONCOCIN ran on the time-shared DEC computers, using a standard terminal for the time-oriented display of patient data. The current version uses compact, single-user workstations running on the SUMEX Ethernet network with large bit-mapped displays for presentation of patient data. The project has also expanded in scope. There are three major research components: 1) ONCOCIN, the therapy planning program and its graphical interface; 2) OPAL, a graphical knowledge entry system for ONCOCIN; and 3) ONYX, a strategic planning program designed to give advice in complex therapy situations. Each of these research components has been split into two parts: continued development of the cancer therapy versions of the system, and generalization of each of the components for use in other areas of medicine. This section will concentrate on the three core research topics derived from our applied work: 1) design of therapy planning systems for use in clinical trial experiments (E-ONCOCIN), 2) implementation of knowledge acquisition systems for clinical trials, and 3) development of general approaches to strategic therapy planning. The work on continued development of the ONCOCIN cancer chemotherapy advisor system itself is described separately in Section IV.A.3.

1 - Overview of the ONCOCIN Therapy Planning System

ONCOCIN is an advanced expert system for clinical oncology. It is designed for use after a diagnosis has been reached, focusing on assisting with the management of cancer patients who are receiving chemotherapy. Because anticancer agents tend to be highly toxic, and because their tumor-killing effects are routinely accompanied by damage to normal cells, the rules for monitoring and adjusting treatment in response to a given patient's course over time tend to be complex and difficult to memorize. ONCOCIN integrates a temporal record of a patient's ongoing treatment with an underlying knowledge base of treatment protocols and rules for adjusting dosage, delaying treatment, aborting cycles, ordering special tests, and similar management details. The program uses such knowledge to help physicians with decisions regarding the management of specific patients.

A major lesson of past work in clinical computing has been the need to develop methods for integrating a system smoothly into the patient-care environment for which it is intended. In the case of ONCOCIN, the goal has been to provide expert consultative advice as a by-product of the patient data management process, thereby avoiding the need for physicians to go out of their way to obtain advice. It is intended that oncologists use ONCOCIN routinely for recording and reviewing patient data on the computer's screen, regardless of whether they feel they need decision-making assistance. This process replaces the conventional recording of data on a paper flowsheet and thus seeks to avoid being perceived as an additive task. In accordance with its knowledge of the patient's chemotherapy protocol, ONCOCIN then provides assistance by suggesting appropriate therapy at the time that the day's treatment is to be recorded on the flowsheet. Physicians maintain control of the decision, however, and can override the computer's recommendation if they wish. ONCOCIN also indicates the appropriate interval until the patient's next treatment and reminds the physician of radiologic and laboratory studies required by the treatment protocol. This core research report begins with our efforts to extend the techniques of ONCOCIN for use in other areas of medicine (E-ONCOCIN).

2 - E-ONCOCIN: Domain Independent Therapy Planning

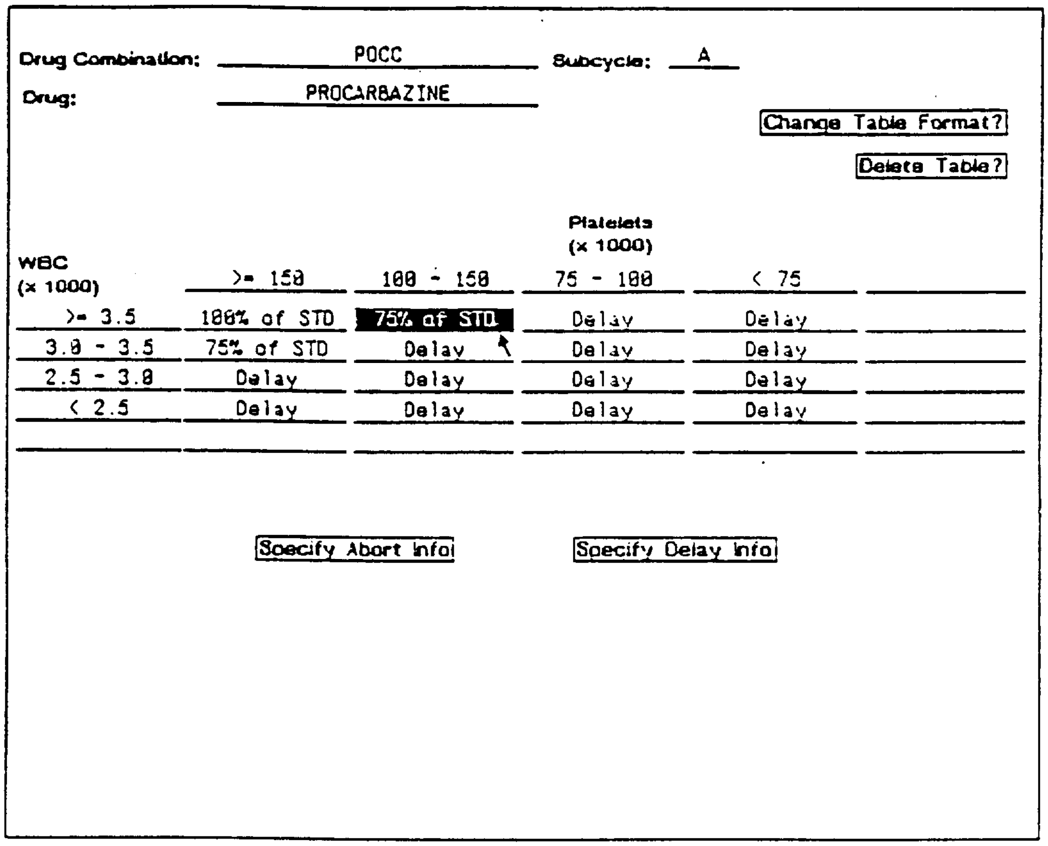
During this past year, our E-ONCOCIN research has concentrated on understanding how protocols in medicine vary across subspecialties. We felt that the area of insulin treatment for diabetes would be a good area to explore. Like cancer chemotherapy, treatments for diabetes continue over long periods of time and have been the area of intensive protocol development. Unlike cancer chemotherapy, the treatment plan must handle multiple doses over the course of one day and deemphasizes the use of drug combinations (although there are a variety of types of insulin). Other challenges of the diabetes area include consideration of multiple goals, such as finding the "normal dose" of insulin versus adjusting for short term trends. Diabetes treatment plans must be flexible enough to take into account diet and exercise patterns and their effects on insulin requirements.

We performed knowledge acquisition sessions about insulin treatment of diabetes, using the medical literature and several internists in the Medical Computer Science research group (Mark Frisse, Mark Musen, and Michael Kahn). The proposed structure for the knowledge base was implemented using the object-oriented programming language upon which ONCOCIN has been based. These experiments, like those of adding more protocols to ONCOCIN, demonstrated the need for changes in the way that the knowledge base can access the time-oriented data base that stores patient data and previous conclusions. The relationships between the different doses and types of insulin treatments will also require alternative ways of building treatment hierarchies. Thus, our initial experiments have shown that many of the elements of the ONCOCIN design are sufficiently general for other application areas, but that some specific elements (particularly the representation of temporal events) will have to be generalized. During the coming year, we will continue our knowledge acquisition experiments and design a version of the E-ONCOCIN system that is separate from the ongoing "clinic version."

3 - OPAL: Graphical Knowledge Acquisition Interface

OPAL is a graphical environment for use by an oncologist who wishes to enter a new chemotherapy protocol for use by ONCOCIN or to edit an existing protocol. Although the system is designed for use by oncologists who have been trained in its use, it does not require an understanding of the internal representations or reasoning strategies used by ONCOCIN. The system may be used in two interactive modes, depending on the type of knowledge to be entered. The first permits the entry of a graphical description of the overall flow of the therapy process. The oncologist manipulates boxes on the screen that stand for various steps in the protocol. The resulting diagram is then translated by OPAL into computer code for use by ONCOCIN. Thus, by drawing a flow chart that describes the protocol schematically, the physician is effectively programming the computer to carry out the procedure appropriately when ONCOCIN is later used to guide the management of a patient enrolled in that protocol.

OPAL's second interactive mode permits the oncologist to describe the details of the individual events specified in the graphical description. For example, the rules for administering a given chemotherapy will vary greatly depending upon the patient's response to earlier doses, intercurrent illnesses and toxicities, hematologic status, etc. Figure 1 shows one of the forms provided by OPAL for this type of specification. It permits the entry of an attenuation schedule for an agent based upon the patient's white count and platelet count at the time of treatment. Tables such as this are generally found in the written version of chemotherapy protocols. Thus, OPAL permits oncologists to enter information using familiar forms displayed on the computer's screen. The contents of such forms are subsequently translated into rules and other knowledge structures for use by ONCOCIN.



Drug Combination: POCC Subcycle: A

Drug: PROCARBAZINE

Change Table Format?

Delete Table?

WBC (x 1000)	Platelets (x 1000) >= 150	100 - 150	75 - 100	< 75	
>= 3.5	100% of STD	75% of STD	Delay	Delay	
3.0 - 3.5	75% of STD	Delay	Delay	Delay	
2.5 - 3.0	Delay	Delay	Delay	Delay	
< 2.5	Delay	Delay	Delay	Delay	

Specify Abort Info

Specify Delay Info

Figure 1: A Sample OPAL Form

Status of the OPAL System

OPAL is one of the few graphical knowledge acquisition systems ever designed for expert systems. Even fewer are designed to be used as the main method for entering knowledge as opposed to a proof of concept implementation. We have pursued three directions in the development of the OPAL system, also in response to the large number of protocols entered through this system during the last year. The first direction is the modification of graphical forms needed to allow the entry of facts that did not show up in the protocols used to test the initial version of OPAL. OPAL continues to assume that most of the knowledge to be entered will have very stereotyped forms, e.g., dose attenuations for most treatment toxicities are based on a comparison of only one laboratory measurement at a time, such as using the BUN to adjust for renal toxicity. We sometimes need much more complex ways of stating the scenarios in which dose adjustments may be necessary. This need has led us in a second direction, towards a "lower-level" rule entry approaching the syntax of the reasoning component of ONCOCIN, but using graphical input devices where applicable. A prototype version of this rule entry system has been completed, and will soon be evaluated as an adjunct to the basic OPAL system.

The OPAL program maps the information provided on the graphical forms into a complex data structure (called the IDS) that is used to represent the contents of the protocol. The data structure is used for copying information from one protocol to another, and as the basis for the creation of the ONCOCIN knowledge base. Our experiments with OPAL, and our intention to generalize OPAL for use outside of oncology protocols, suggested that we reorganize the OPAL program to use a relational database to store its knowledge. We have patterned the database after an existing database query syntax. Because no relational database management systems exist for the Interlisp language upon which OPAL is based, we reimplemented the database from its written description. The database structure is now almost complete, and we have begun to design a revised IDS for chemotherapy protocols, and will be determining how an IDS would be created for other areas of medicine (e.g., the insulin example being used in the E-ONCOCIN experiments).

Our ability to use the OPAL system for specifying oncology treatments has led us to the design of a new program, named PROTEGE, that will turn an interactive session with an expert and knowledge engineer into the specification of an OPAL-like system for clinical trials in a wide range of medical areas. We have implemented several prototype forms for PROTEGE. These forms are used to specify a general description of the application area. Of particular importance is the need to specify how the therapy planning process will take place, e.g., how will the initial dosage of a drug be combined with various adjustments of the dosage due to toxicities to the treatment to form the final recommended dose. Most of this type of "procedural" knowledge is not entered in the OPAL system, and must be hand-coded by the knowledge engineer. A Ph.D. thesis on PROTEGE is in progress by Mark Musen, M.D., and will be completed during the next year.

4 - ONYX: Strategic Therapy Planning

Although the knowledge of cancer chemotherapy is rich and complex, protocols seldom refer directly to underlying models of drug action. The guidelines in a protocol are, rather, high-level composite descriptions of expert advice, based on the study designers' experience as well as biological models of the therapeutic agents and their mechanisms of action. We have observed, however, that when protocols fail to cover a complex clinical situation that arises for a given patient, expert oncologists will turn to underlying mechanistic models and use them to assist in the decision-making process. ONCOCIN has no such knowledge; it must therefore occasionally decline to make a recommendation and instead refer a physician to the study chairman for a decision about how to manage a particular complex problem. It is accordingly a long-range goal to add model-based expert-level reasoning to ONCOCIN's performance.

Our research in model-based reasoning is embodied in a program known as ONYX. This system is based on the observation that creative planning strategies in the oncology domain (and many other fields) appear to involve a three-step process: (1) heuristic generation of a small number of plans, i.e., plausible responses to the problem at hand, (2) mental simulation (also called "envisionment") of how the patient would respond over time if each of those plans were carried out, and (3) selection of a preferred plan based upon the likelihood of the various possible outcomes *and* the value placed on those outcomes by the patient and physician. Step 2 in this process involves patient-specific simulation of tumor pathophysiology and drug action, but it also depends on recognition that the outcomes of interventions cannot be predicted with certainty and that probabilistic predictions are more realistic. Thus, model-based probabilistic simulations in ONYX are coupled to a decision analytic module which assists with the third step in the process. The work outlined here is preliminary.

Each of the components in ONYX may be generalized for use in other systems. We have concentrated our work on the decision analysis component. We are building tools that will allow experts to frame the comparison between several possible treatments that could be administered at one point in a patient's course. Often these treatments will be variations on the standard treatment, but with reduced dosages or delayed time of treatment. An important part of the treatment decision concerns the patient's evaluation of the possible outcomes and their likelihood, as represented in the *utility* of the various plans. The program we have built carries out a dialog with patient to assess the utilities, builds a decision tree, and prints out the "best" choice. A graphical representation of the decision problem is build on the computer display as the dialog takes place.

A major problem with decision analysis programs have been the way that the choice is explained to the user. Often, the answer is in the form of one utility number for each choice. Most computer systems for decision trees allow the user to see how much the utilities will change as the probabilities of the expected events are modified. What is not available, is an explanation, in English, of why one choice is better than another. As part of his Ph.D. research, Curtis Langlotz has built a system that can create a rationale for the selection. The program compares various parts of the decision tree, looking for differences in the problem structure that account for the variation in the final utilities for the problem. This explanation program has been tested with several decision problems from different areas of medicine: treatment of heart disease, antibiotic selection, and cancer treatment.

5 - Implementation of the ONCOCIN Workstation in the Stanford Clinic

In mid-1986, we placed the workstation version of ONCOCIN into the Oncology Day Care clinic. This version is a completely different program from the version of ONCOCIN that was available in the clinic from 1981-1985 -- using protocols entered through the OPAL program, with a new graphical data entry interface, and revised knowledge representation and reasoning component. One person in the clinic (Andy Zelenetz) became primarily responsible for making sure that our design goals for this version of ONCOCIN were met. His suggestions included the addition of key protocols and the ability to have the program be useful for clinicians as a data management tool if the complete treatment protocol had not yet been entered into the system. Both of these suggestions were carried out during this year, and the program has achieved wider use in the clinic setting. In addition, laser-printed flowsheets and progress notes have been added to the clinic system.

The process of entering a large number of treatment protocols in a short period of time led to other research topics including: design of an automated system for producing meaningful test cases for each knowledge base, modification of the design of the time-oriented database and the methods for accessing the database, and the development of methods for graphically viewing multiple protocols that are combined into one large knowledge base. These research efforts will continue into the next year. In addition, some of the treatment regimens developed for the original mainframe version are still in use and can be transferred to the new version of ONCOCIN. The process of converting this knowledge will also be undertaken in the next year. As the knowledge base grows, additional mechanisms will be needed for the incremental update and retraction of protocols.

We also developed new insights about the design of the internal structures of the knowledge base (e.g., the relationship between the way we refer to chemotherapies, drugs, and treatment visits). We will continue to optimize the question-asking procedure, improve the method for traversing the plan structure in the knowledge base, and consider alternative arrangements used to represent the structure of chemotherapy plans. Although we have concentrated our review of the ONCOCIN design primarily on the data provided by additional protocols, we know that non-cancer therapy problems may also raise similar issues. The E-ONCOCIN effort is designed to produce a domain-independent therapy planning system that includes the lessons learned from our oncology research.

6 - Personnel

The development of the generalized version of each of the ONCOCIN components has been undertaken by a large group of computer scientists and physicians. Samson Tu has had primary responsibility for the extensions to the design of the knowledge base, Clifford Wulfman has had primary responsibility for extensions to the data entry interface. David Combs has had primary responsibility for the knowledge acquisition interface. Janice Rohn has been involved with protocol and data management, and has primary responsibility for the implementation of the program that sets up the ONCOCIN user environment. Christopher Lane has developed the object-oriented systems software upon which the entire ONCOCIN system is designed.

III.A.3.3. Core AI Research

1 - Rationale

Artificial Intelligence (AI) methods are particularly appropriate for aiding in the management and application of knowledge because they apply to information represented symbolically, as well as numerically, and to reasoning with judgmental rules as well as logical ones. They 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.

Expert systems are one important class of applications of AI to complex problems -- in medicine, science, engineering, and elsewhere. An expert system is one whose performance level rivals that of an human expert because it has extensive domain knowledge (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, BB1, and 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.

There is considerable power in the current stock of techniques, as exemplified by the rate of transfer of ideas from the research laboratory to commercial practice. But we also believe that today's technology needs to be augmented to deal with the complexity of medical information processing.

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 Knowledge Systems Laboratory (KSL), 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 contribute in the long term to improved knowledge based systems in biomedicine.

2 - Highlights of Progress

In the last year, research has progressed on several fundamental issues of AI. As in the past, our research methodology is experimental; we believe it is most fruitful at this stage of AI research to raise questions, examine issues, and test hypotheses in the context of specific problems, such as management of patients with Hodgkin's disease. Thus, within the KSL 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. Many of these specific problem domains are medical or biological. In this way we believe the KSL 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 KSL technical reports listed in the publications section.

2.1 - 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 continues on BB1, with its explicit representation of control knowledge, as reported last year (see the summary of Blackboard Architectures below). In addition, part of our research on NEOMYCIN is focused on using a flexible, rich representation of control knowledge so that we can model problem solving at the strategic level as well as at the tactical level.

[See KSL technical reports KSL-87-01 and KSL-87-32]

2.2 - Blackboard Architectures and Control

How can we design flexible control structures for powerful problem solving programs?

We have continued to develop the BB1 blackboard architecture for systems that reason about -- control, explain, and learn about -- their own actions. In the area of control, we have developed two new domain-independent control capabilities. One generic control knowledge source refines specified parameters of abstract control plans by generating legal values from a semantic network. The other control knowledge source performs opportunistic goal-directed reasoning whenever actions recommended by other control decisions are not executable. In the area of explanation, we have developed the ExAct program. It provides a flexible, menu-driven set of explanation alternatives, as well as a graphical display of the comparative advantages of alternative actions. In the area of learning, we have developed two new capabilities. The WATCH program observes domain experts solving problems and attempts to abstract from their actions the underlying control strategy. It automatically programs new control knowledge sources to generate the hypothesized strategy on subsequent problems. The TRANALOGY program notices when problems in a new domain are analogous to problems in a known domain. It hypothesizes that analogous reasoning methods will work in the new domain as well and automatically programs appropriate knowledge sources.

We have begun conducting various experiments on the costs and benefits of control reasoning. In the context of the PROTEAN system for protein structure modeling, we are investigating the power of different kinds of control knowledge and strategies to produce computational efficiency. Early results suggest that a small computational investment in control reasoning can produce substantial computational savings in problem-solving operations. We also are exploring differences among alternative architectural realizations of a particular control strategy.

We have continued to develop the ACCORD framework for the class of arrangement problems exemplified by PROTEAN: arrange a set of objects to satisfy constraints. ACCORD substantially enhances BB1's general capabilities for control, explanation, and learning. In addition to PROTEAN, we have applied BB1-ACCORD in the SIGHTPLAN system for designing construction site layouts.

In order to accommodate ACCORD and other task-specific frameworks, we have developed a set of generic framework interpretation procedures for: parsing framework sentences, matching and rating sentences, generating legal parameter values for sentences, and translating sentences into the lower-level language of BB1. These procedures apply to any user-specified framework that satisfies the standards of knowledge and representation laid down in ACCORD. We refer to this growing collection of systems and knowledge modules as the BB* environment.

[See KSL technical reports KSL-86-38, KSL-87-8, and KSL-87-10 and "other outside publications" in Section III.A.3.5]

2.3 - Advanced Architectures

The goals and technical approach of this project, largely supported by DARPA under the Strategic Computing Program, have been discussed in previous annual reports. To summarize briefly, we seek to achieve two to three orders of magnitude speedup in the execution of knowledge-based systems, by identifying and exploiting sources of concurrency at all levels of system design: the application level, the problem solving framework level, the programming language level and the hardware systems architecture level. Due to the inherent complexity of the task and the lack of theoretical foundations for parallel computation with ill-structured problems, we have taken an empirical approach. During the first phase of the project, which will be concluded in July, 1987, we have made specific choices at each of the system levels, i.e. taken a "vertical slice" through the design space, and have conducted several experiments to investigate the effects of a wide variety of parameters on performance.

Some highlights of our accomplishments thus far (most of which occurred during the past year) include:

- Based on a careful and systematic study of potential hardware system architectures, we have established an architectural framework for the underlying machine as a multicomputer array. The study ranged over the full spectrum of possibilities, from shared memory multiprocessors to shared memory multicomputer networks to distributed memory multicomputer networks, taking into account the VLSI opportunities of the 1990's.

- We have designed and constructed a complex, fully instrumented simulator to realize the above architectural framework. The simulated class of machines, called CARE, permits full manipulation of the parameters which specify the hardware system, e.g. communication topology, memory size, etc. CARE is written in Zetalisp, and runs on standard Lisp workstations (TI Explorer, Symbolics 36xx).

- We have studied and implemented basic additions to the Lisp language to accomplish distributed Lisp processing on CARE class machines. These additions are now incorporated into the basic simulation language.

- We created an initial, experimental operating system for CARE class machines, called CAOS. CAOS was used to produce our first experimental results, an end-to-end experiment using the ELINT application, using replicated knowledge sources and pipelining for achieving parallel activity.

- The results of these early experiments were encouraging. Linear speedup, close to the 45 degree line, was achieved up to the intrinsic limits of the application.

- We generalized the traditional blackboard problem solving concept, and developed two new blackboard frameworks. These two frameworks, CAGE and POLIGON, take opposite points of view with respect to the locus of computing activity. CAGE uses knowledge sources as the active agents, whereas POLIGON takes a view that is oriented more towards dataflow, in which the blackboard nodes are the active agents.

- We evaluated a variety of real-world applications as drivers of the underlying system levels, discarding several candidates which initially looked promising but turned out not to be, for various reasons. Consequently, we decided to build our own application, AIRTRAC. As we programmed this application in different problem solving frameworks we began to learn techniques for parallel programming. We initiated experiments to study the performance of AIRTRAC in both blackboard frameworks.

- Detailed studies of the performance achieved in the ELINT/CAOS experiments led to drastic simplification of the pipelining scheme, an orientation toward implementing blackboard nodes as active agents, and using parallel object oriented programming as a low level implementation technique. An environment, called LAMINA, grew out of this analysis. Experiments are in progress to compare the performance of AIRTRAC implemented in LAMINA with AIRTRAC implemented in the blackboard frameworks. The first set of AIRTRAC/LAMINA experiments, using part of the knowledge base that can be used in a data-driven manner, exhibited linear speedup close to the limit of the concurrency inherent in the task.

By the end of 1987 we will have completed five sets of vertical slice experiments. It is already clear that these experiments could have significant impacts on both the hardware and software communities. Specifically:

- One important impact of our research will be to shift the emphasis in parallel architectures for knowledge-based systems from (probably premature) building of hardware to the development of software systems, techniques and tools for the encoding of knowledge-based applications. Hardware can certainly be built. The real difficulty is in developing a firm, quantitative understanding of what hardware actually matters and what hardware may actually hurt (e.g., building hardware based upon incompletely thought-out policy decisions in the software design).

- We will have demonstrated that the distributed memory paradigm is not only a viable alternative to shared memory architectures, but perhaps superior in important ways. The vertical slice experiments provide evidence that implementing a relatively complex application, using a non-shared address space with message passing, can be accomplished without the complexities of managing shared address spaces. Moreover, we will have demonstrated that distributed-memory multicomputers can be programmed to achieve significant (ten to one hundred times) speed-up for nontrivial symbolic problem solving applications. Furthermore, such multicomputer systems will provide a better fit to the (forecasted) technology for ULSI of the 1990's than the shared memory architectures.

- We will have demonstrated that the major "source of power" in parallel computing is the ability to allow the user to express and manipulate parallel constructs at the level of the application. Thus, the best return on investment is to develop appropriate tools to support parallelism at this level, rather than to support the development of the underlying languages or compilers. The speedup obtainable by only parallelizing programming language constructs in a "programmer transparent" manner (e.g., parallel Prolog or parallel production systems) is very limited.

- An important lesson learned from the success of our simulator is that real applications can be carefully analyzed in an instrumented environment, thereby permitting experimentation with alternate architectures. The community would do well to stress simulation over hardware building.

- We will have demonstrated the need for fast process creation and process switching mechanisms.

[See KSL technical memos KSL-86-36, KSL-86-69, KSL-87-02, KSL-87-07, KSL-87-34, KSL-87-35.]

2.4 - Knowledge Acquisition and Machine Learning

Our research in machine learning has focused on several distinct problem domains including medical (NEOMYCIN/HERACLES) and biochemical (PROTEAN) in addition to domain-independent investigations. We also are motivated by the need for effective tools for knowledge acquisition and maintenance of knowledge bases (IMPULSE and STROBE for FRM, BBEDIT, KSEDIT with BB1).

Several papers by researchers in the KSL were presented at AAAI-86 in Philadelphia in August. Wilkins and Buchanan describe a method of debugging rule sets (see below). Rosenbloom and Laird [14] present a mapping between the SOAR architecture and explanation-based generalization (EBG), in which a justifiable concept definition is acquired from a single training example and an underlying theory of how the example is an instance of the concept. SOAR is an architecture that supports general learning through chunking, which is similar to but not the same as EBG. In addition, the authors suggest answers to some of the outstanding issues in explanation-based generalization.

Chunking is a learning mechanism that acquires rules from goal-based experience. SOAR is a general problem-solving architecture with a rule-based memory that can use the learning capabilities of chunking for the acquisition and use of macro-operators. Rosenbloom et al. are investigating chunking in SOAR and find that chunking obtains extra scope and generality from its intimate connection with the sophisticated problem solver (SOAR) and the memory organization of the production system.

In their AAAI-86 paper, Horvitz, Heckerman, and Langlotz present a framework for comparing alternate formalisms for plausible reasoning [6]. They demonstrate a logical relationship between several intuitive properties for measures of belief and the axioms of probability and discuss its relevance to research on reasoning under uncertainty in artificial intelligence.

Inductive Rule Learning

Buchanan, et al. present an empirical study of the incremental learning process using a careful selection of counter examples in concept formation with the rule-learning system RL (described in last year's SUMEX report). They find that "near misses", negative examples that are similar to acceptable cases, are particularly effective in shrinking the space of possible theories that explain the examples observed. They define a metric for the distance of each example from the target theory and measure the effectiveness and efficiency of examples related to the distance measured, demonstrating that the power of near misses to restrict the space of possible theories results from their small distance from the target. They also find that intelligent selection of instances based upon knowledge of the state of the evolving theory results in a faster convergence of an evolving theory toward the target concept, requiring many fewer cases for learning.

Debugging Knowledge Structures

In large rule-based systems, the performance of the system is strongly dependent on the degree to which the knowledge of the system is "debugged" and refined, i.e., erroneous rules are identified and removed, redundant rules are combined, missing rules are added, and certainty factors of rules are found that give good results over many cases. Such evaluation and restructuring of knowledge is an important type of learning and can be automated to some extent. Here we describe recent work in the debugging and refinement of knowledge bases using several techniques.

Wilkins and Buchanan [19] analyze a problem with the rule sets of rule-based systems that use certainty factors, i.e., better individual rules do not necessarily lead to a better overall set of rules. Since all less-than-certain rules contribute evidence towards erroneous conclusions for some problem instances, the distribution of these erroneous conclusions is not necessarily related to the quality of individual rules. This has important consequences for automatic machine learning of rules, since rule selection is usually based on measures of quality of individual rules. The authors present a method using a new Antidote Algorithm that performs a model-directed search of the rule space to find an improved rule set. They report that the application of this method significantly reduces the number of misdiagnoses when applied to a rule set generated from 104 training instances. This work was also presented at the AAAI-86 Conference in August.

Debugging the knowledge structures of a problem solving agent is the *synthetic agent* method [20] determines a *performance upper bound* for debugging a knowledge base. The synthetic agent systematically explores the space of near miss training instances and expresses the limits of debugging in terms of the knowledge representation and control language constructs of the expert system. This paper presents the framework for evaluating a differential modeling system.

Wilkins describes the ODYSSEUS apprenticeship learning program [21], designed to refine and debug knowledge bases for the HERACLES expert system shell. ODYSSEUS analyzes the behavior of a human specialist using two underlying domain theories, a *strategy theory* for the problem solving method (heuristic classification), and an *inductive theory* based on past problem solving sessions. ODYSSEUS improves the knowledge base for the expert system shell, identifying bugs in the system's knowledge in the process of following the line-of-reasoning of an expert, serving as a knowledge acquisition subsystem. ODYSSEUS can also be used as part of an intelligent tutor, identifying problems in a novice's understanding and serving as student modeler for tutoring systems.

Wilkins, et al. illustrate that an explicit representation of the problem solving method and underlying theories of the problem domain provide a powerful basis for automating learning for expert system shells [22]. By using domain-independent *task procedures* and *task procedure metarules*, domain knowledge can be located and applied to achieve problem solving subgoals. However, these rules are often limited in use due to insufficient domain knowledge. This paper describes the use of *metarule critics* in ODYSSEUS for automating the acquisition of domain knowledge, illustrating a powerful form of failure-driven learning at the level of subgoals as well as at the level of solving the entire problem.

III.A.3.4. Core System Development

1 - Introduction

In this section we describe progress on our core system development and work toward a distributed AIM community. Before launching into the technical details, the motivations and plans for core system work are first summarized along four dimensions: 1) the motivation for the shift of the SUMEX-AIM community from a central mainframe-based model of computing resources to a largely distributed workstation-based model; 2) the prospects for workstation technology and vendor support for a diverse distributed AIM community; 3) the core SUMEX-AIM systems tasks needed to complement vendor developments to realize distributed community operation; and 4) the integration, dissemination, and management of the shift of the AIM community from a centralized to a more distributed operation, including the remaining central resource functions:

- *Motivation for a Distributed Resource:* The motivations for supporting and managing the AIM community as a distributed community are manifest. First the cost/performance trade-offs between centralized shared computing facilities and personal workstations have shifted dramatically toward workstations, especially in the area of interactive symbolic computation resources. While the technology is still quite young, the very best environments for developing knowledge-based systems for biomedicine are arguably already on personal workstations. Various kinds of workstations are rapidly decreasing in cost and increasing in performance so that appropriate models can be selected for cost-effective research support or system dissemination into practical settings like health care clinics or application laboratories.

  Second, the AIM community, with its growing ties into other diverse areas of biomedical informatics, has long been too large to effectively support from a single central node like SUMEX. A number of AIM groups have already moved to local mainframe computing resources (such as at Rutgers University, the University of Pittsburgh, the University of California at Santa Cruz, the University of Minnesota, and Ohio State University). Only some of these have been able to establish network connections for their machines to date, without which low-speed terminal connections must still be made to the central SUMEX resource for mail exchange, software sharing, information access. As workstation prices fall, this trend toward decentralization will accelerate and the need for uniform network access, information services, and systems/software support will increase. The challenge will be to provide responsive central resource services that encourage and facilitate effective communication, collaboration, and information sharing in the new distributed environment.

- *Prospects for Workstation Technology:* Computer workstations have already demonstrated remarkably high performance and low cost for symbolic computing applications. The prospects for future generations of workstations promise an even fuller spectrum of price/performance alternatives. Even with the trend toward more effective personal workstations, however, there are still aspects of an overall computing environment most effectively implemented and supported through central resources. These include services like large-volume information and file storage, special parallel computing architectures, multi-vendor systems expertise, and experimentation with integrating new computing technologies for community deployment. But hardware is only a small part of the picture -- software represents the larger challenge in the effective integration of workstations with shared resources -- and here is where a community systems integration effort is required. Most vendors are motivated to maximize the sales of their own products, whereas a community of the size and scope of the AIM community must be prepared to integrate technologies from diverse vendors in order to maximize its productivity and to keep abreast of rapidly developing new capabilities. The role of SUMEX-AIM in this new era is to integrate what is available from diverse vendors with core system development efforts to facilitate community research and communications and the smooth evolution of the AIM distributed computing environment.

- *Core Systems Development Tasks:* In order for workstations to support AIM community activities with minimum dependence on expensive, central mainframes, they must be able to supply not only outstanding knowledge-based system development environments but also general computing environments for tasks like electronic communications, text processing, information and file management, and utilities like spreadsheet systems. Many workstation environments do not have fully developed facilities in all these areas and must be augmented. Another major area of core system effort will be in the development of tools to facilitate effective workstation to workstation interactions. These tools include being able to access remote workstation and central computing resources, linking the graphics displays of remote workstations with each other over communication networks, establishing and managing cooperative computing tasks, and enabling remote transfer and sharing of files and information. Finally we must stay abreast of the rapidly changing workstation technology and have allocated a small amount of funding each year to purchase appropriate examples of systems important to AIM community research for testing, evaluation, and development.

- *Managing the Community Transition:* As system research and development progresses, much will remain to be done to integrate and disseminate these new workstation tools throughout the national AIM community -- so that the central DEC 2060 resource can be phased out while maintaining support of community activities. System tools must be tested, evaluated, and refined in the broad context of the AIM community; community groups must fund, acquire, install, and learn to use suitable workstation and network communications equipment; residual central services must be developed and made accessible to support sharing software tools, user consulting, and information resources; and AIM workshop and other management tools for coordinating, integrating, and extending community activities must be evolved. We will use a small group of Stanford and AIM community AI researchers and students to guide the development and testing of distributed subsystems throughout the research period. Initially, these will come mainly from the Stanford community which is easily accessible and has a long experience in experimenting with the development and use of workstation technologies for AI research. After the early years of development and experimental dissemination, we will begin to introduce these tools more extensively for general AIM community use. Our estimate is that these tasks will require the full five-year research period in order to carry out the necessary development, make an orderly and smooth transition, and evaluate the results, without disrupting communications or inter-group collaborations.

2 - Remote Workstation Access, Virtual Graphics, and Windows

2.1 - Remote Access

Lisp workstations of various types have proven extremely powerful, both as development environments for artificial intelligence research and as vehicles for disseminating AI systems into user communities. In addition to the compact, inexpensive computing resources workstations provide, high-quality graphics play a key role in their power. Such graphics systems have become indispensable for understanding the complex data structures involved in developing and debugging large AI systems and are important in facilitating user access to working programs (e.g., for ONCOCIN and PROTEAN). However, as we move towards a distributed workstation computing environment for AI research in the SUMEX-AIM community (and move away from the centralized, shared DEC 2060), a number of technical obstacles must be overcome. One of the most important is to eliminate the need for the user display to be situated close to the workstation computing engine.

This is important in order to allow users to work on workstations over networks from any location -- at work, at home, or across the country. The first step has been getting reliable terminal access operational on all workstations. All workstations now have TCP/IP based terminal servers, and TCP/IP is being installed in the SUMEX network terminal concentrators. This allows primitive (non-graphical) access to the workstation's abilities. A more comprehensive access will be provided through our remote graphics work.

2.2 - Virtual Graphics

In the past, members of the SUMEX-AIM community have often watched each others programs work by linking their CRT terminals to the text output of a running program on the SUMEX 2060. In the case of workstations, though, it is much more difficult to link across several networks to view the complex graphics output of a program. Even locally, it is important to make graphical interaction with workstations across campus or from home possible. One would like to be able to provide the same powerful graphical tools and programming environment that are available to a user sitting in front of the workstation to the remote user if that user has a low-cost bit-mapped display and mouse. In order to accomplish this, it is necessary to capture and encode the many graphics operations involved so that they can be sent over a relatively low-speed network connection with the same interactive facility as if one had the display connected through the dedicated high-speed (30 Mhz) native vendor display/workstation connection.

As reported last year, we studied the feasibility of remote access to workstations by experimenting with a virtual graphics protocol, the Virtual Graphics Terminal Service (VGTS), which was developed at Stanford in the Computer Science distributed systems group [9, 8]. The VGTS provides tools to define objects like windows, lines, rectangles, circles, bitmaps, ellipses, splines, and graphics events like mouse clicks independently of the graphics hardware and operating systems. This encoding minimizes the communication bandwidth required between cooperating hosts, to remotely draw a line, for example.

We also reported that an implementation of this protocol was developed and installed in the operating system of a Xerox 1186 Lisp workstation so that its presence would be transparent to the programmer. This means that if one connects to such a LISP workstation from a SUN workstation (running suitable VGTS software), the Lisp machine graphics will be sent over the net and reconstructed on the SUN workstation without changes to the application program running. This implementation has worked very well in early experiments so that over an Ethernet, the remote response time is quite close to the response time on the Lisp machine itself.

As a consequence of this work, we had demonstrated the feasibility of remotely using LISP workstations over an Ethernet to take advantage of their graphics programming environment.

During the past year, two new contenders for a virtual graphics standard protocol appeared. These were the MIT Project Athena X window system [15], and Sun Microsystems, Inc.'s Network Extensible Window System [17], referred to as *X* and *NeWS*, respectively. We spent several months studying both X and NeWS and met with representatives of each group supporting these protocols.

X is a very complete protocol that has been developed over the past several years at MIT[1]. X operates at a somewhat lower level than VGP, and as a result can be more bandwidth-intensive. It also assumes a static allocation of computation, display, and interaction responsibilities between server and client. On the other hand, it more fully implements the event mechanisms necessary to track mouse/window interactions and mouse motion histories, and supports color. The protocol has been quite carefully thought out, and provides more flexibility for implementing reasonable emulations of the variety of window systems that exist within our environment. For example, TI Explorers have mouse-sensitive regions within windows called "active regions," and X allows the support for such a region by defining an *Input Only* window with its own cursor. When the mouse moves into such a window, the cursor changes to show the user that he has entered an active region, and at the same time sends an enter-window event to the client. The client can then take the appropriate action for that active region (for instance, scroll text). This is impossible to do in VGP.

NeWS is unique in the sense that it uses a programming language to define its protocol. This programming language is an extension of Adobe's PostScript page layout language for laser printers. This feature gives NeWS its extensibility, for if one wishes to add a new function to the server, one simply sends the PostScript procedure implementing it to the server, and remotely executes that new procedure. This gives the client a great deal more control over what a window looks like; for example, one could implement round or elliptical windows with NeWS. NeWS also allows a client to interact with mouse motion histories and mouse/window events. Thus, it was very difficult to choose between these two protocols.

Ultimately, we chose X as the remote graphics protocol standard for our work. This decision was pragmatic, since we have limited staff resources, and X is receiving wide support from both vendors and the Common Lisp community. An X *client* implementation is being written for Texas Instruments Explorers here at SUMEX-AIM[2]. Our TI Explorer X client is well underway. It is being written in Common Lisp and uses *flavors*, the Explorer object system, to represent instances of X windows. We are currently beta-testing Xerox Common Lisp, and will port the Explorer X client to our Xerox Lisp Machines later this year.

Currently, TI in conjunction with MIT is developing a server implementation for Explorers. DEC is a major supporter of X, and there are implementations under development for their Vax line of equipment. Sun Microsystems is also doing an X

---

[1]The X protocol has been completely redefined this past year. Its most recent version, X.11, is assumed in all of the discussion that follows.

[2]The *client* software runs on the Lisp machine and sends the graphics protocol commands to the remote user display system. The dual of the client is the X *server* software which runs on the user display system and translates the X protocol sent by a client Lisp machine into real graphics pictures and mouse actions.

implementation beneath NeWS, as well as porting X to run directly on their equipment. We are an alpha test site for the SUN implementation. This will provide us with preproduction X server software that we can run on our SUN workstations to aid in debugging our own client software. We anticipate implementations for workstations like MacIntosh II's when a production version of X is released this Fall.

The X window protocol is more bandwidth intensive than some other protocols. It is our feeling that even with this limitation, a suitable subset of the X protocol can be used in cross-country connections where slower communications speeds and longer delays are common. We will have to determine empirically what this subset is. One, for example, would not want to track a mouse in such a situation, but could reasonably expect to use mouse/window events, such as *EnterWindow* or *LeaveWindow*, to manage a remote display over long connection distances. In any case, more work needs to be done in this area to fully develop and integrate these capabilities into Lisp machine systems and to insure that cross-country connections will indeed give usable response time. Success of this work will mean that one can use LISP machine systems from TELENET, ARPANET, or an Ether TIP connection throughout the SUMEX-AIM community.

2.3 - Remote Graphics Applications

As an example of applying the remote graphics ideas, a TALK program has been implemented which facilitates interactive, electronic communication between users on independent workstations. Layered on the workstation's native editor, the program allows the full use of all editing capabilities in the process of communication, including deletions, corrections and insertions, font changes, underlining, paragraph formatting, etc. Since the workstation's editor also supports both low- and high-level graphics, the program not only facilitates textual exchanges among users, but also allows the sending of screen images (back traces of program breaks, code fragments, etc.) as well as structured graphics images (which can be modified on the destination workstation and returned), all interactively. An example of a TALK session and an illustration of TALK's relationship to other subsystems in the workstation software environment are shown in Figure 2.

The TALK program allows the use of different user interfaces, the workstation's document editor being just one possibility. We also implemented a simpler terminal mode for compatibility with similar programs on other similar and dissimilar workstations. The program was implemented initially using the Xerox XNS family of Ethernet protocols for convenience and speed of development to try out the ideas. Future extensions will include allowing use of different Ethernet (and possibly non-Ethernet) protocols, since the program only requires a reliable byte-stream to operate. We expect the IP/TCP protocols will be added next in order to be able to use the program over the ARPA network.

The TALK program was released gradually to increasing numbers of users in order to get real users' feedback and make changes accordingly. The Medical Computer Science group did an extensive test of the system, where for a period, they used it in place of their normal electronic and non-electronic communication methods whenever possible. This was both a test of the program and an exploration into what people want in the next generation of electronic communication. The TALK program has been released to the Xerox Lisp workstation community as a whole and researchers at Xerox PARC successfully used the program to hold an interactive, graphic, electronic conversation between users at the PARC facility (in California) and Xerox's EuroPARC in England.

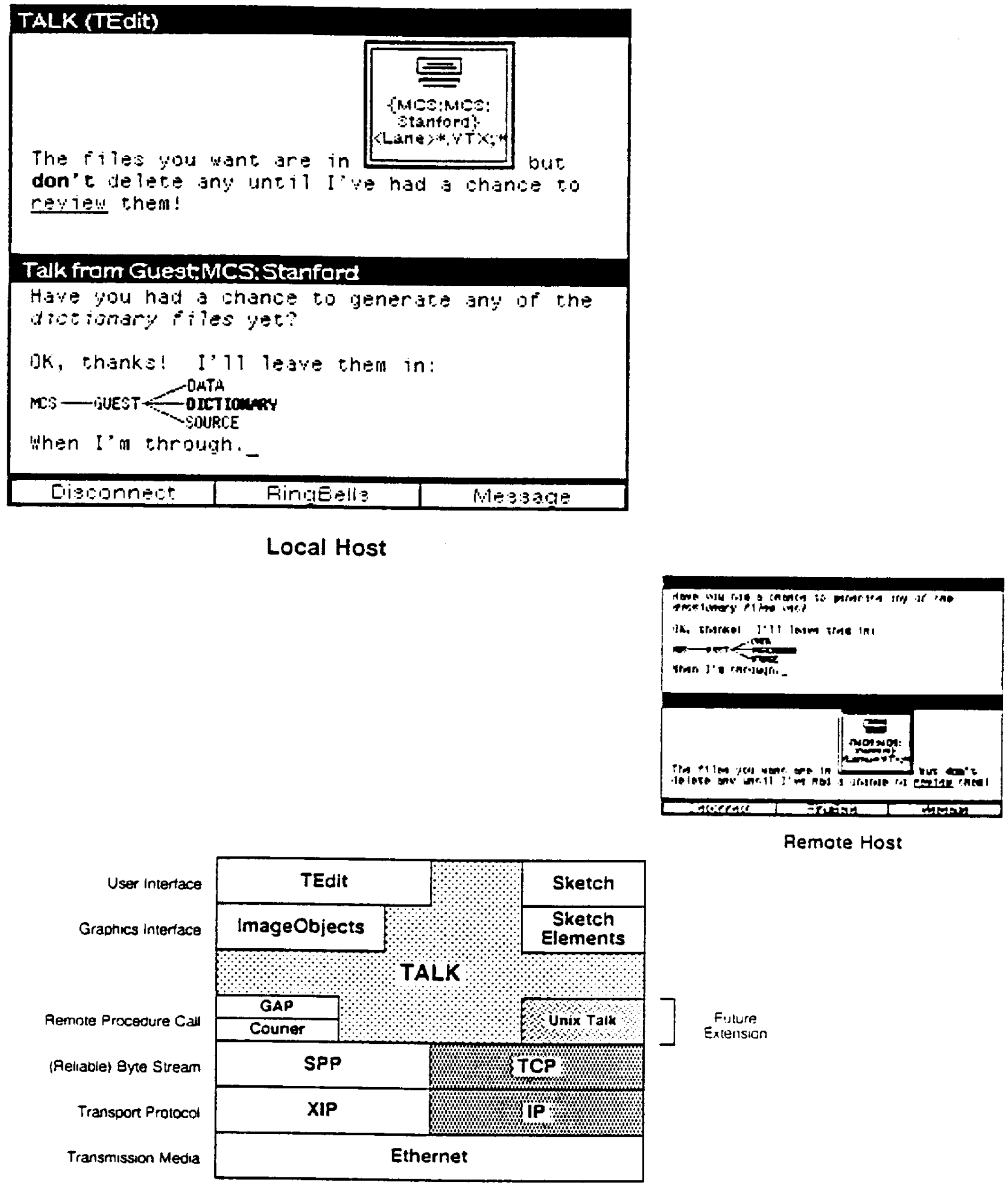


Figure 2: TALK Session Example and the Software Layers Involved in TALK

2.4 - Application-level Window System Standards

Modern programs need to utilize the multiple presentations, non-textual images, and non-keyboard inputs available on all the systems in use by SUMEX. However, up until now, each machine's window system has been idiosyncratic to that machine. There is considerable research now aimed at providing a powerful, flexible window system that can be implemented on a wide variety of hardware, and utilized by many forms of software. However, most of this research is directed at the primitive operations needed to do basic graphics, windowing, and interaction (as in the discussion of X protocols above). We are also working to develop a high level interface to a standard windowing system targeted at the writer of AI applications programs. This system is not being designed to specify the entire man/machine interface, but to provide a simple, easy to understand and useful way for program authors to provide sophisticated interfaces without spending a large percentage of their time working only on the interface. We are currently in the midst of analyzing current applications in order to develop a model for this system based on real-world experience.

3 - File Access and Management

A stable, efficient mechanism for storing and organizing data is central to any computing environment, and is one of the most challenging issues in the move to distributed, workstation-based computing. It is necessary to provide standard services, such as file backup, archival, a flexible, intuitive naming facility, and data interchange services (e.g., software distribution). We also feel that, as the amount of data being manipulated grows, it will become more and more important to have powerful tools for managing hierarchies of files. We plan to support the community with a number of UNIX-based file servers, like the VAX-based servers in use at SUMEX for several years (see Figure 7) and the new SUN-based server (see Figure 5). These will require continued SUMEX-AIM development, however. By keeping the number of servers small, the distributed namespace problem should be manageable in the near term. Current UNIX file servers are relatively cheap and fast. UNIX has many of the needed facilities, e.g., backup, long names, hierarchical directory structure, some file property attributes, data conversion, and limited archival tools. However, while general issues of networking, remote memory paging services, and flexible file access have received considerable attention in both the academic and commercial development of file servers, there seems to be little attention given to other critical operational needs. For instance, the much-used file archiving system of the DEC 2060 (sometimes called off-line cataloged storage) has no analog service in the UNIX systems. Perhaps this is the result of UNIX having its origin in the small computer world where the number of users and volume of data has traditionally been quite low. Our efforts are going into improving the archival facilities and providing case independence and multiple generations by adding SUMEX software between the file system and the network. This should temporarily solve these problems without substantial loss of performance or maintainability.

For the long-term use of the distributed community, we plan to develop an optical disk-based backup and archival system and to use enhanced tools on workstations to do file management. We are currently investigating hardware options for optical disk systems. As better techniques for managing a distributed file system come out of the early research stages, we will use them to improve the distributed file service facilities.

3.1 - Remote File Access

During the past year, there has been a welcomed progress in vendors' attempts to standardize file access protocols. Previously, each vendor had addressed the file storage needs of their particular workstation in a way that was incompatible with most other workstations, making shared file access and support difficult in a highly heterogeneous environment such as the SUMEX-AIM community. Also, the resources required to maintain many distinct families of filing conventions and protocols on specialized hardware, all meeting the performance needs of a demanding research community, are prohibitive. Thus, last year we proposed to adopt a variant of the *NFILE* file access protocol[1] developed by Symbolics, Inc. It now appears, however, that Sun Microsystems, Inc.'s (SMI) Network File System (NFS) is becoming a more prevalent industry standard, despite the fact that it does not support extensible file attributes and file generations. In order to encourage the porting of NFS to other vendors' workstations, SMI has placed NFS in the public domain, and has a special group dedicated to aiding interested parties in writing the requisite software. This group is also willing to make some changes to the protocol to support non-UNIX file systems (for example, they recently made a change so that NFS could be ported to a CRAY computer). We are now beta-testing a Texas Instruments implementation of NFS on our Explorers, and are ourselves engaged in implementing NFS on Xerox Lisp workstations.

Given that we have acquired an experimental SUN file server this year, and that NFS is supported in the Kernel of the 4.3 release of Berkeley UNIX, this path for unified file access across our mix of workstations appears to be the best solution available. Our anticipated move to 4.3 UNIX on our VAX file servers this summer, and the completion of the NFS port to the Xerox Lisp machines will give us a single file access protocol that is supported by all of our systems with the exception of the Symbolics 3600's. It appears that a third party is working on an NFS implementation for Symbolics machines and we will test this in the coming year.

3.2 - File Server Throughput

At present, a number of file service strategies are employed among and within the various workstation and time-sharing communities. Each strategy has its merits and drawbacks and only in their aggregate do they address all the needs of the users.

One yardstick of utility is the maximum speed of data transfer. Speed of data transfer is affected by the speeds of the processors, disks, I/O circuitry, file system design, network transport protocols, file service protocols, software efficiency, system loading, and other operational parameters. Simple throughput measurements suggest that for the immediate future, the mixed-vendor file service strategy still has advantages from the point of view of data transfer speed. (See Figure 3.)

For the Xerox workstations, the Xerox 8037 file server (using the NS Filing protocol) provided the greatest measured throughput (roughly 37% faster than the Sun 3/180 and Vax 11/750 file servers, using TCP FTP). For the TI workstations, the fastest server was another TI Explorer (using the Chaos FILE protocol) providing throughput 91% greater than the nearest contender (a vax using the Chaos FILE protocol), and 269% faster than the closest IP/TCP contender. The Sun workstation provides a virtual file system interface only for the Sun NFS protocol, and hence was not benchmarked against alternative servers because we are still working on optimized NFS facilities for other workstations and servers.

None of the client/server configurations tested approached the theoretical maximum throughputs projected by disk speeds, network speeds, and other system design considerations. Therefore, we believe that through more effective software engineering it will be possible to simultaneously improve data transfer speed and to reduce the number of server implementations necessary to support the present level of service. For example, the potential for software improvement was illustrated this year by fine-

[1] A *file access protocol* is intermediate between a *remote file system* and a *file transfer protocol.*

tuning of the Xerox implementation of TCP, which yielded improved Sun file server throughput by a factor of 30. In the immediate future, our experiments in this area will focus on the new implementations of the *NFS* file service protocol.

Client	*Server*	*Protocol*	*Reading Throughput*	
DEC 2060	Sun 3/180	TCP FTP	7,000 baud (loaded)	
Sun 3/180	DEC 2060	TCP FTP	17,000 baud (loaded)	
Sun 3/75	Sun 3/180	TCP FTP	55,000 baud (unloaded)	
Xerox 1186	DEC 2060	PUP Leaf	18,181 baud (loaded)	
Xerox 1186	DEC 780 (VMS)	TCP FTP	33,402 baud (loaded)	
Xerox 1186	Xerox IFS	PUP Leaf	52,526 baud (unloaded)	
Xerox 1186	DEC 750 (UNIX)	PUP Leaf	53,036 baud (loaded)	
Xerox 1186	DEC 2060	PUP FTP	67,001 baud (loaded)	
Xerox 1186	Sun 3/180	TCP FTP	71,192 baud (unloaded)	(was 2,412 baud)
Xerox 1186	DEC 2060	TCP FTP	72,207 baud (loaded)	(was 2,850 baud)
Xerox 1186	DEC 750 (UNIX)	TCP FTP	72,412 baud (loaded)	(was 9,096 baud)
Xerox 1186	Xerox IFS	PUP FTP	84,125 baud (unloaded)	
Xerox 1186	Xerox 8037	NS Filing	103,519 baud (loaded)	
Xerox 1186	Xerox 8033	NS Filing	105,486 baud (unloaded)	
Xerox 1132	DEC 2060	TCP FTP	3,228 baud (loaded)	
Xerox 1132	DEC 2060	PUP Leaf	18,737 baud (loaded)	
Xerox 1132	DEC 750 (UNIX)	PUP FTP	75,361 baud (loaded)	
Xerox 1132	DEC 750 (UNIX)	PUP Leaf	81,711 baud (loaded)	
Xerox 1132	DEC 750 (UNIX)	TCP FTP	121,163 baud (loaded)	
Xerox 1132	DEC 2060	PUP FTP	167,687 baud (loaded)	
Xerox 1132	Sun 3/180	TCP FTP	215,000 baud (loaded)	
Xerox 1132	Xerox 8037	NS Filing	234,154 baud (loaded)	

			Reading Throughput	*Writing Throughput*
TI Explorer	DEC 750 (UNIX)	TCP FTP	36,952 baud	96,000 baud
TI Explorer	Sun 3/180	TCP FTP	58,888 baud	135,208 baud
TI Explorer	TI Explorer	TCP FTP	61,376 baud	121,512 baud
TI Explorer	DEC 2060	TCP FTP	63,320 baud	110,592 baud
TI Explorer	DEC 750 (UNIX)	Chaos FILE	122,136 baud	129,376 baud
TI Explorer	TI Explorer	Chaos FILE	233,008 baud	221,192 baud

Figure 3: File Server Throughput Benchmarks

4 - Electronic Mail

Electronic mail has become a primary means of communication for the widely spread SUMEX-AIM community. The advent of distributed workstations is forcing a significant rethinking of the mechanisms employed to manage such mail. With mainframes, each user tends to receive and process mail at the computer he uses most of the time, his *primary host*. The first inclination of many users when an independent workstation is placed in front of them is to begin receiving mail at the workstation, and, in fact, many vendors have implemented facilities to do this. However, this approach has several disadvantages:

- Workstations (especially Lisp workstations) have a software design that gives full control of all aspects of the system to the user at the console. As a result, background tasks, like receiving mail, could well be kept from running for long periods of time either because the user is asking to use all of the machine's resources, or because, in the course of working, the user has (perhaps accidentally) manipulated the environment in such a way as to prevent mail reception. This could lead to repeated failed delivery attempts by outside agents.

- The hardware failure of a single workstation could keep its user "off the air" for a considerable time, since repair of individual workstation units might be delayed. Given the growing number of workstations spread throughout office environments, quick repair would not be assured, whereas a centralized mainframe is generally repaired very soon after failure.

- It is more difficult to keep track of mailing addresses when each person is associated with a distinct machine. Consider the difficulty in keeping track of a large number of postal addresses or phone numbers, particularly if there was no single address or phone number for an organization though which you could reach any person in that organization. Traditionally, electronic mail on the ARPANET involved remembering a name and one of several "hosts" (machines) whose name reflected the organization in which the individual worked. This was suitable at a time when most organizations had only one central "host." It is less satisfactory today unless the concept of a "host" is changed to refer to an organizational entity and not a particular machine.

- It is very difficult to keep a multitude of heterogeneous workstations working properly with complex mailing protocols, making it difficult to move forward as progress is made in electronic communication and as new standards emerge. Each system has to worry about receiving incoming mail, routing and delivering outgoing mail, formatting, storing, and providing for the stability of mailboxes over a variety of possible filing and mailing protocols.

Thus, we are investigating the alternative strategy of having a *mail server* machine which handles *mail transactions*. Because this machine would be isolated from direct user manipulation, it could achieve high software reliability easily, and, as a shared resource, it could achieve high hardware reliability, perhaps through redundancy. The mail server could be used from arbitrary locations, allowing users to read mail across campus, town, or country using more and more commonly available workstations.

The mail server acts as an interface among *users*, *data storage*, and *other mailers*. Users employ a *mail access protocol* (MAP) to retrieve messages, access and change properties of messages, manage mailboxes, and send mail. This protocol should be simple enough to implement on relatively uncomplicated, inexpensive machines so that mail can be easily read remotely. This is distinct from some previous approaches since the mail access protocol is used for all message manipulations, isolating the user from all knowledge of how the data storage is used. This means the the mail server can utilize the data storage in whatever way is most efficient to organize the mail. The data storage could be anything from conventional magnetic disk file system to a highly specialized mail filing system built on optical disks, since it is abstracted from other elements in the mail system. The other mailers constitute the mail server's (and thus the users') link to the outside world. The mail server would use various *mail transport protocols* (e.g., SMTP) to exchange mail with other mail hosts.

We have been investigating user/mail interface issues for workstations, as well as issues for the mail access protocol itself. We are examining several related projects, including MIT's PCMAIL (Mark Lambert, MIT Distributed Systems Group), the public parts of Xerox's Grapevine and NSMail, and work on Stanford's V system.

We have implemented an Interim Mail Access Protocol (IMAP) server on the 2060 and a client implementation in Interlisp on Xerox D-machines. The resulting beta-test mail environment proved to be quite usable; some D-machine users use it as an alternative to the 2060 mail environment in their daily mail work.

The IMAP server manipulates the actual file store copy of the user's incoming electronic mail under direction from the IMAP client. As noted above, the client has no knowledge of the (possibly operating system- dependent) format of mail on the server's file store; the IMAP protocol provides its own representation of mail and the server translates between this and its host system file store conventions.

The IMAP client issues a series of *fetch* commands to retrieve data from the server. A fetch command has two arguments: a *message sequence* and the name of the data item to be fetched. A message sequence can be a single message number, a range of message numbers, or a list of numbers or ranges. For example, a typical fetch command might be "fetch 2:7,10 flags", meaning "fetch the status flags for messages 2 through 7 and message 10" (status flags include "new message", "deleted message", "message has been read", etc. as well as user-defined flags).

In IMAP, the actual message data is identified by names such as "RFC822.Header" and "RFC822.Body" referring to the text-based mail representation used on the DoD Internet standard (RFC 822). This is intended to be a temporary solution only, since RFC 822 lacks structure and the capability to deal with non-text mail. We plan on extensions to IMAP (IMAP II, see below) that will introduce a canonical and structured representation of an electronic mail message. In such a structured form, an electronic mail message would consist of a set of named properties and property values.

During implementation of the user interface we observed that the IMAP protocol had several deficiencies which made certain mail concepts difficult or impossible to implement. For example, there is no way in IMAP to notify the client of newly arrived mail during an IMAP session. Other IMAP deficiencies were observed in the design of a Common Lisp implementation for Texas Instruments Explorers; in particular, IMAP is a "lock-step" protocol with no mechanism for multiplexed operation. This means that IMAP is vulnerable to synchronization problems in which a client interprets part of a previous response as the answer to the current query.

To address these concerns, a new Interim Mail Access Protocol (IMAP II) was designed after extensive review. IMAP II is heavily influenced by IMAP, although with a greater degree of formality in the specification and quite a bit more extensibility. Instead of the lock-step query/response model of IMAP, IMAP II uses tagged commands and data and explicitly allows unsolicited data to be sent from the server to the client. IMAP II introduces a more formal structure to server-to-client path; all data is now identified unambiguously. This is especially important for extensibility and unsolicited data.

In addition, IMAP II makes it possible to fetch more than one item of data at a time. This is an important performance issue since often the client needs to fetch a set of data items for a set of messages. The IMAP model of fetching a single data item at a time resulted in the client having to make several consecutive requests with much longer turnaround than a single request that specifies everything the client wants.

A large subset of IMAP II has been implemented on the 2060 by modifying the existing IMAP implementation. Both the 2060 implementation and the specification have been left open-ended to allow for extensions as the need arises. Work is now in progress to modify the Interlisp user interface to use IMAP II. Since the interface is no longer limited to the model of IMAP a general restructuring of the Interlisp client is being done to take advantage of the new facilities offered by IMAP II. A Common Lisp implementation, based on IMAP II, is also in progress.

5 - Text Editing

All workstation systems have text editing facilities, some adaptations of systems in use on mainframes (e.g., EMACS-like editors) and some specialized What-You-See-Is-What-You-Get (WYSIWYG) editors (e.g., TEdit for Xerox workstations or InterLeaf for UNIX workstations). We are currently making use of each workstation's facilities, making extensions where needed to bring compatibility among the various workstations (in both user interface and document format) without detracting from the powerful, but idiosyncratic, features. Text formatting, to produce printable or displayable forms of documents, is another area where considerable vendor effort is expended. Implementations of SCRIBE or TEX systems are available for some workstations directly. Also, since these formatting processes are essentially batch operations, we expect to provide servers that offer formatting services. These can be fed a raw manuscript and will return a formatted version, suitable for one of the several printing device standards in use. WYSIWYG editors are able to combine the editing and formatting processes into the document preparation system. We will concentrate on PostScript and ImPress printers, allowing Press printers to fade from use. The 2060 also provides for printer spooling, based on a first-come-first-serve algorithm with priorities determined by submission time and estimated pages of output. This spooling is not available among workstations currently. Given adequate printing resources, a laissez-faire access policy without spooling can work adequately. If there is a problem, an arbitration scheme will need to be worked out, but this should be a relatively straightforward task. Finally, we will need to augment vendor products to provide essential text processing aids for functions like spelling correction, document merging and segmenting, and document analysis.

5.1 - Text Processing for Xerox D-machines

TEdit is the text editing and formatting package on the Xerox D-machines (i.e., the 11xx series) and we have continued our work to extend this environment to displace text processing from the DEC 2060. Almost all efforts during the past year were directed towards the Interlisp package TMAX. TMAX stands for Tedit Macros And eXtensions and it gives TEdit the ability to do things that hitherto could only be done with Scribe. Scribe is a powerful document preparation language, but it consumes far too many mainframe cycles. Furthermore, with Scribe you must hardcopy your output to see what it looks like. TEdit is a WYSIWYG text editing and formatting system, which means that you can see what your output will look like while you are creating it.

TMAX makes no attempt to mimic Scribe in TEdit. This would be a Herculean task given the power and flexibility of Scribe. Instead TMAX implements some of the more commonly used features of Scribe, including indexing, numbering, end notes, and forward/backward referencing. All of these features are implemented through menus.

For example, to include an index request in a document, the user simply buttons the Index command (with the mouse) and then types the text to be indexed. TMAX takes care of all the rest (e.g., associating the page number with the indexed text, creating a sorted list of the indices, etc.). These TMAX features plus the editing and formatting features already available in TEdit make the TMAX/TEdit package an attractive alternative to Scribe.

The following is a quick overview of the major TMAX features:

- Indexing -- users can insert both simple and extended index requests, create a sorted file of the indices and their page numbers, and even specify that the page numbers be printed in manual format (e.g. III:25.7 for chapter 3, section 25, page 7). A simple index is just the text to index. An extended index takes the text to sort on, the text to print, the font to print it in, and a page number option. This option allows the user to specify the normal page number in the index file, no page number, or a user specified fixed page number. There is also a command that pops up a menu of the simple and extended indices specified so far and users can insert additional index requests by simply buttoning the corresponding item in this menu.

- Numbering -- users specify the names and order of "number markers" and then insert these markers wherever they want something numbered. Users can create as many different number markers as they like and some can be layered (i.e. chapter, section, etc.) while others are disjoint. When a marker is inserted or deleted, TMAX automatically adjusts all the related numbers. Users can even specify the font and format of each number. The format defines how the number will be displayed (i.e. an Arabic or Roman numeral or a letter), the delimiter following the number, and the starting value. There is also a facility to create a standard table-of-contents file.

- End Notes -- these are just like footnotes except end notes are inserted at the end of the text rather than the bottom of the page. A future version of TMAX will support footnotes. When an end note is inserted or deleted, TMAX automatically adjusts the other end note numbers.

- References -- users can refer to specific numbering markers or end note numbers by their numeric value. It does not matter if the number is before or after the reference to it. Also, should a number change because a number marker or end note was deleted or inserted, the reference to that number will be automatically updated (as well as the number itself).

There are many more features and options in TMAX and still more in the planning stages. For example, one can edit the text of an end note by pointing the mouse to the end note number and pressing the middle button. Another TEdit window will appear containing the end note text. Some of the features planned are footnotes, bibliographies, and appendices. The TMAX User's Guide describes all the features of this package.

5.2 - Remote Editing

Currently, the mainframe editor of choice among our users is EMACS. EMACS, like Scribe, is very powerful but it also places a heavy load on our mainframe. In an effort to reduce the mainframe load (and ease users into using TEdit), we have written WEDIT (Workstation EDITor). WEDIT provides a convenient way for mainframe EMACS users to edit their files on a Xerox D-machine using TEdit. Note that WEDIT itself is not an editor. It simply opens a connection to the workstation and sends a packet containing the name of the file to edit. The workstation does all the rest. When the user is done editing, the workstation sends the updated file back to the mainframe. From the mainframe's point of view WEDIT is an editor in the sense that given a file, it returns an updated version. Because of this, it is easy to install as the default mainframe editor. A simple change in the user's login command file is all that is necessary. From that point on, each time the user edits a file, the editing will be done by TEdit on the user's personal workstation. EMACS users can experience long delays when the 2060 is heavily loaded. With WEDIT (i.e. TEdit) there are no delays since the editing is done on the user's personal workstation.

5.3 - Special Document Types

In last year's report, some TEdit extensions to facilitate simple document types like memos were mentioned. These extensions proved to be very useful although this package was only a prototype. Using the concepts developed in this package, we have written a new TEdit package called *Letterhead*. The Letterhead package allows users to create standard letterheads, for example, for Stanford University correspondence. All the options in this Letterhead package are menu driven. When a user starts the Letterhead package, a TEdit window appears on his workstation and the user is prompted for several different fields. First a menu of the possible Stanford logos pops up and the user must select one of these logos. The logo is placed in the upper left hand corner of the window. Next a menu of the return addresses pops up. The user may select one of the known addresses or create his own. Next the Letterhead package asks the user how the address should be justified. This is done though a menu and the possible ways are left, right, or centered. The justified address is then inserted in the upper right hand corner of the window. Finally, the current date is automatically inserted just below the logo. Now the letterhead is complete and TEdit is ready to accept input from the user. The user can change either the logo, address, or date by pointing the mouse at the appropriate field and pressing the middle button. If the logo is buttoned, the logo menu will pop up and the user can select a different logo. If the address is buttoned, the return address menu will pop up and the user can either select a known address, create his own address, or edit the address already in the document. If the date is buttoned, a date menu pops up allowing the user to display the date in one of several different formats. When this window is hardcopied, it will look just like a standard Stanford University letter

6 - System Building Tools

Traditionally, a large set of languages and programming environments have been supported on the 2060 in order to encourage experimentation and development. We now believe that the experience gained in those years of broad experimentation can be distilled into a fairly small set of languages and tools, relieving the researcher of the need to learn many programming languages, while still providing the needed facilities to allow the experimentation to move further into the higher levels of knowledge representation systems and problem solving architectures. As we move to the workstation based environment, we plan to phase out support for many of the languages we have offered in the past and concentrate on the most relevant languages for AI research and applications: C, FORTRAN, InterLisp-D, ZetaLisp, and Common Lisp. Common Lisp has already achieved popularity as a standard (see page 54), and many projects are already using it. We expect to press for further adoption of Common Lisp as a community symbolic computing standard, consistent with prior investments in large software systems such as those which exist for on-going AIM projects. In addition, we will support important higher-level knowledge representation and problem solving architectures (e.g., S.1, KEE, Strobe, and others) as appropriate for community research and dissemination activities.

7 - Distributed Information Resources and Access

There are many user needs for getting information from and about the computing environment, ranging from help with command syntax to sophisticated database queries. A distributed computing environment adds new complexities in making such information accessible and also new requirements for information about the distributed environment itself. We are adapting the many workstation-specific information tools to include distributed environment information such as workstation and server availability, "finger" information about user locations and system loads, network connectivity, and other information of interest to users in designing approaches to carrying our their research tasks. In addition, we will have to develop general systems tools for monitoring and debugging distributed system performance to identify workstation and network problems. Finally, we must adapt and develop distributed system tools for remote database queries and organize the diverse sources of information of interest to AIM community members to facilitate remote workstation access to community, project, and personal information that has traditionally existed in ad hoc files on mainframe systems.

In conjunction with the SUN file server we have been integrating, we have mounted an experimental database system for remote information access using the commercial UNIFY database product. Our goal is to make access to the database information possible from a distributed workstation environment through network query transactions, as opposed to asking the user to log into the database system as a separate job and type in queries directly. This will facilitate remote information access from within *programs*, including expert systems, where the information can be filtered, integrated with other information, and presented to the user. The system will provide multi-user, multi-database access capability; that is, several users will be able to have access to a single database at the same time, and a single user will be able to have access to several databases at the same time.

The initial implementation of the remote query system was done on a TI Explorer. The query interface on the Explorer communicates with the Sun UNIFY database system via the Remote Procedure Call (RPC) mechanism which underlays the NFS remote file access system. The Explorer calls a server on the SUN and sends an SQL/DML query command as an argument to a remote query procedure, and receives the retrieved data and/or a message sent back from the server. SUN UNIFY can already manage multiple databases, so a client can have several databases open at the same time. The operations on the database are transaction-oriented, and therefore the concept of a database access session is applicable. The access functions currently implemented are *open a database*, *close a database*, *retrieve data from a database*, *insert records into a database*, *delete records from a database*, *update the database*, *lock a database*, and *unlock a database*.

This facility can be easily converted to run in Lisp environments on machines with SUN RPC services implemented. Currently, there is no RPC package for the Xerox D-machines, so we undertook implementing one. This should be done by early summer.

8 - Distributed system operation and management

The primary requirements in this area are user accounting (including authorization and billing), data backup, resource allocations (including disk space, console time, printing access, CPU time, etc.), and maintenance of community data bases about users and projects. Our accounting needs are a function of NIH reporting and cost recovery requirements. The distributed environment presents additional problems for tracking resource usage and will require developing protocols for recording various kinds of usage in central data base logs and programs for analyzing and extracting appropriate reports and billing information. We are now involved in analyzing the kinds of resource usage that can be reasonably accounted for in a distributed environment (e.g., printing, file storage, network usage, console time, processor usage, server access), and investigating what facilities vendors have provided for keeping such accounts. Data backup is, of course, closely related to the filing issue. We continue to use and improve network based file backup for many of our file servers.

9 - Mainframe and Workstation System Environments

The various parts of the SUMEX-AIM computing environment require development and support of the operating systems that provide the interface between user software and the raw computing capacity. This includes the mainframe systems and the workstation systems. Following are some highlights of recent system software environment developments.

9.1 - TOPS-20

Our long-term plan to phase out the 2060 mainframe system has continued as scheduled. Development efforts on the 2060 have ceased, except where needed to keep the machine operational in the evolving distributed environment. This involves considerable work in areas such as file system archiving, retrieval, and backups; periodic updating, checkout, and installation of new versions of system software; the regular maintenance and updating of system host and network tables; and monitoring of and recovery from system failures, both hardware and software. Over the past year, the main areas of activity include:

- *Network service reliability* -- The SUMEX 2060 has experienced relatively frequent software crashes resulting from system problems in the handling of free space by the IP/TCP network software. During periods of heavy use, the entire system would suffer an unscheduled restart approximately every eighteen hours. After a considerable amount of investigation of crash dumps, we isolated a cause. The problem was introduced over a year before in a modification made by another site in an attempt to improve network performance. After fixing this illusive bug, the 2060 reliability has improved markedly and the system regularly runs for over a week between reloads.

- *Network naming domains* -- The Internet community is in the process of converting to a domain naming scheme, to replace the flat address space of the old exhaustive host tables prepared by the Network Information Center. Although we have converted to using only fully qualified names, we are not yet running the domain system on the 2060. This is due in part to the unreliability and incompleteness of the domain software for TOPS-20's at this point. We expect to move to full domain support this coming year.

- *Dial-up communications* -- A significant portion of work on the 2060 is carried on via dialup modems from homes. During the past year we rearranged and consolidated our incoming modem lines. We combined several inside and outside phone number hunting sequences serving several different modem types and speeds, into well defined groups for old-style Vadic 1200 modems, local versions of split speed modems, and other types. This last group serves any Bell/CCITT modem at any speed from 300 baud to 2400 baud. During this process we removed all the outside phone lines, and now operate exclusively through Stanford-operated SL100 lines. In addition to these mainframe modems, we have installed 10 modems on an Ethernet TIP, allowing users, once dialed in, to connect to the host of their choice.

- *Cost Center accounting* -- During the past year, the 2060 accounting programs were updated to reflect the new Cost Center structure (see Section III.D.2). All the various users and projects were organized according to their cost center account numbers, and monthly reports are generated to reflect this usage. As part of this conversion, a concerted effort was made to review all of the SUMEX accounts, and remove those that were otherwise no longer appropriate.

9.2 - UNIX

We run UNIX on our shared VAX 11/750 file servers. This system has been used pretty much as distributed by the University of California at Berkeley, except for local network support modifications, such as for ChaosNet protocols. The local VAX user community is small, so we have not expended much system effort beyond staying current with operating system releases and with useful UNIX community developments.

9.3 - Xerox D-Machines

Much of the SUMEX-AIM community continues to use InterLisp, including many Dandelion (1108), Dandetiger (1109), and DayBreak (1186) machines, in addition to the Dorado (1132). We have used the Xerox implementation of the TCP network protocol (in cooperation with Xerox) extensively this past year and saw its performance and reliability improve a great deal. We began a Lisp implementation of Sun NFS (*Network File System*). The ARPA protocol suite, which is seeing increasing usage, lacks a mechanism for random file access or attribute manipulation. The Sun specification partially fills this void and appears to be a standard whose acceptance is growing.

The Interlisp software remained stable this year and almost no user time was wasted on software revision problems. A number of new utilities were written locally or acquired from other sites with whom we exchange expertise on the ARPA Internet.

We are among the first users of Xerox Common Lisp for the Xerox Lisp machines. The advantages to our community are early availability of this widely-recognized dialect of the Lisp language and the ability to specially direct the implementers' attention to the problems of greatest concern to us.

The Info-1100 discussion list which we sponsor saw another year of growth of readership and participation on the ARPA Internet, Usenet, Bitnet, and CSNet. Among the beneficiaries are other NIH-sponsored projects at Ohio State University and the University of Maryland.

In conjunction with the Info-1100 mailing list, a library of user-written software is made available to the Internet community on the SUMEX-AIM 2060 computer. Over 60 packages and supplements were distributed this way. Additionally, the source code to many of these packages was mailed to the Info-1100 mailing list in order to reach an even wider group.

We have worked closely with many other sites, including the Center for Study of Language and Information at Stanford, the Stanford Campus Networking group, Rutgers University, Ohio State University, the University of Pittsburgh, Cornell, Maryland, and industrial research groups such as Xerox Palo Alto Research Center, SRI, Teknowledge, IntelliCorp, and Schlumberger-Doll Research. We have been the maintainers for the international electronic mail network of users for research D-machines, which have upwards of 300 readers, and the interchange of ideas and problems among this group has been of great service to all users.

Although numerous Xerox Lisp machine sites are able to obtain software from SUMEX-AIM via anonymous FTP over the ARPA network, it became increasingly clear that a large part of the community did not have such access even though there is electronic mail connectivity. To experiment with distributing software to these sites, we put together a simple ASCII encoder for binary files, BMENCODE. This program makes it possible to mail binary files (TEdit editor files and *COM files from the compiler) to isolated sites, exploiting Interlisp's inherent ability to encode bitmaps into ASCII files. Numerous files were successfully transferred around using this program. As the user community has begun to see the value of such a utility, more efficient versions of the program have been developed elsewhere.

In extending our XNS boot service (which provides installation and diagnostic programs for our workstations) to work with the new 1186 Hardware, we ran into trouble as the 1186's hastily written initial Ethernet microcode. The booting sequence violated Ethernet layering principles which prevented it from routing beyond the local network. After nearly a year of exchanging letters, packet traces and software with Xerox, the problem is still unresolved. This led to our adding a second Xerox 8000-based XNS boot server (using a spare 1108 processor) to our other major network with 1186 hardware. This additional server provided a suitable work-around to the problem and only a single workstation is still unable to access network boot services.

Our move to a new building this past year involved the de-installation and reinstallation of nearly thirty workstations plus several printers and other servers. In anticipation of the move, diagnostics were run on all of the Xerox University Grant 1108s in order to get any existing problems fixed under warranty. The diagnostics were run again after the machines were installed in the new facility. All the equipment was successfully relocated without major incident.

9.4 - Texas Instruments Explorers

The twenty Texas Instruments Explorers have enjoyed an increasing popularity as more projects have developed a need for the combination of execution speed, full Common Lisp, and sophisticated development facilities offered by the Explorer. Explorers have come into use in other parts of the national biomedical community as well, such as Ohio State University and the University of Maryland. However, the Explorer is still maturing as an AI workstation. Thus, our efforts have been directed at improving the environment of the Explorer by developing software, organizing user interest activities, and advising Texas Instruments.

Previous experience has shown that the greatest source of advancement for a particular computing environment is the user community. They are the most in touch with the deficiencies of the system, and thus uniquely positioned to address them, as well as to utilize the strengths of the system. The product developers of the system are frequently too involved in the lower levels of detail to produce general, effective solutions to problems, as well as being hampered by limited manpower resources. However, a significant amount of time and effort is required to organize this effort. This task has traditionally fallen to a user-run organization, such as DECUS or Usenix.

We are spearheading the effort to organize a national or international users' group for the Explorer. The goals of this undertaking are to:

- facilitate dissemination of information by organizing meetings where presentations and discussions can be used to make little-known techniques and facilities more widely known, as well as feeding back information on needs and wants to developers,
- allow more immediate communication via electronic mailing lists, which are

used for distribution of important software fixes and discussion of items of general interest, such as new software tools, or proposed changes to the system,

- publish a periodic newsletter containing usage tips, salient extracts from the electronic mailing lists, and announcements,

- and, perhaps most importantly, establish and maintain a library of public domain, user supplied software.

A preliminary meeting was held at AAAI '86, and a second meeting is being planned for AAAI '87. Over 80% of those who have expressed interest in the users' group are members of the Info-TI-Explorer and Bug-TI-Explorer mailing lists, currently maintained on the SUMEX-AIM 2060. Negotiations with Texas Instruments over the legal ramifications of the user library are in the final stages. The format and procedures of the library have been mapped out, and are currently undergoing peer review. Online copies of the library will be maintained at Texas Instruments, and on the SUMEX-AIM 2060, to facilitate ARPANET access to the software.

There are already many entries ready for the library, most of which have been developed locally. We have maintained the software tools that were produced previously by fixing bugs, making improvements, and porting to new releases. Some of these have remained essentially the same, including:

- The Symbolics 36xx to Explorer compatibility package
- The Source Code Controller (was known as the System Manager)
- Imagen Via TCP (was Net Imagen)
- Finger Via TCP (was TCP Finger)
- Vertically Ordered Menu Columns
- General Named Structure Message Handler
- DEFSTRUCT Type Checking
- Batch Processor
- Choice Facility Enhancements (was Choose Variable Values Macros)
- Backup To File System (was FS To FS Backup)

Many of the tools have been enhanced or newly written this year, including:

- A number of pieces that allow the user to exploit a "desk top" usage metaphor, where several applications can be active, or semi-active at once, with the interaction area, or "window" of each application potentially overlapping others. These pieces include:
    - WINDOW-MANAGER-SYSTEM-MENU provides a replacement to the standard system menu that allows for easy manipulation of the placement and shape of windows on the screen, as well as other common display management operations.
    - RUBBER-BAND-RECTANGLES which allows easy, precise specification of a rectangle on the screen by providing a constantly

updated "ghost" image of the rectangle (like a rubber band attached to the mouse), as well as the ability to change corners, and specify a minimum size. Previously, this was done by placing only the upper left and lower right corners, with no ghost box, and only a beep to indicate that the box was too small.

- BACKGROUNDS providing a "curtain" between windows representing active applications, and temporarily inactive ones. In the desk-top metaphor, this adds a drawer to the desk. A menu of background operations, as well as eye-pleasing images, are also provided.

- DEEXPOSED-MOUSE which allows windows to handle mouse clicks and documentation even when they are not completely exposed.

- SNAPSHOT-WINDOWS which allow the user to copy a portion of the screen, thus saving it for later use.

- TRANSPARENT-WINDOW which allows the image under a window to bleed through, providing the illusion of non-rectangular windows.

- An on screen round analog clock, with sweep second hand.

- Development tool consistency enhancements, including:

  - commands in the debugger, data structure inspectors (regular and flavor), editor so that they can call each other

  - commands in tools that do not have them to call the Lisp evaluator, obtain argument lists, obtain macro expansions, call the compiler, trace function invocation, and obtain programmer supplied documentation.

  - showing editor buffer reading status in a fashion similar to file reading status at the bottom of the screen

  - the ability to call the debugger on stack groups from many different contexts

  - the ability to modify entries in inspect panes in applications other than the Inspector

- Facilities to display a graph of time-varying quantities. This facility is useful for monitoring system performance parameters, such as the number of network packets sent or received, the number of disk operations per second, or the amount storage allocated.

- A Screen Saver that shuts the display video off after about twenty minutes of keyboard idleness to reduce display phosphor deterioration.

- A version of the terminal emulator program that does not take up the entire screen, and can have user configurable fonts.

- A facility for attaching functions to arbitrary keyboard keys, most commonly used to cause a particular instantiation of an application to be selected when a key is pressed, allowing rapid movement among applications.

- A number of editor commands, including Tags Compile Macro Calls, Macro Expand Into Buffer, Rotate Buffer, Rotate Buffer Backwards, Add File To Tag Table, Remove File From Tag Table, and Evaluate And Insert Into Buffer.

- Functions that allow the user to map over a set of files, applying a function to each file. The set of files can be specified very generally, and values accumulated in various ways from the mapped function.
- Extensions to the flavor inspector allowing it to function in many situations where it would previously have failed.
- A tools for displaying data organized hierarchically in trees, or as graphs, featuring
    - full cycle detection and handling,
    - mouse sensitive nodes and edges,
    - dynamic editing of the graph display,
    - horizontal and vertical scrolling,
    - an "overview" mode to facilitate moving the view port around in a large graph
- Introductory documents which have been used as models by a number of sites.

Of course, all of these will be provided to the user's library, and many of them have already been given to other sites, including Intellicorp, Berkeley, ISI, University of Maryland, and Ohio State.

In addition to producing and maintaining these software tools, we attempt to provide extensive testing and evaluation of Explorer hardware and software products in a sophisticated university research environment in order that these products work more effectively when they are distributed to the national community. This testing is critical to the development of the computing environment since the combination of concentrated in-house expertise and close links to the product developers allows a turnaround on problem fixes unavailable in the broader scope.

This year we have participated in testing TI's implementation of the Network File System protocol, Release 3.0 of the Explorer System Software, and Release 2.0 of TCP/IP. Our testing of NFS, besides uncovering the usual set of bugs, has allowed us to make suggestions to TI that have led to an order of magnitude increase in the data throughput of the implementation. Similarly, our experience with DARPA Internet protocols has allowed us to make many suggestions for improving the Release 3.0 Namespace System which TI has claimed to be invaluable in making the system acceptable to the many Arpanet users in the national community.

We served as a test site for several hardware revisions, as well, and further plan to perform extensive testing of the Explorer II VLSI-based machine when it becomes available in late spring or early summer.

Third party software is less utilized, but we stay abreast of the latest releases of the expert system shell KEE, and will be evaluating the Scribe text formatting system on the Explorer in a matter of weeks.

In addition to specific testing and evaluation, we are constantly finding, tracking, fixing, and reporting software bugs. This year we submitted thirty-two new bug reports on Release 2.1, twenty-one of which had fixes included. All of these fixes have been made available to the national community in a patch file.

There have been fifteen formal reports, with ten fixes on release 3 beta in the current four weeks of testing. There were forty-two reports returned with the TI representative who brought the initial software after the first week, most of which have been fixed by TI.

We have also worked extensively on the operational issues involved in keeping the machines running and useful from day to day. Texas Instruments had no hardware or software maintenance plans for large university installations in place, so we worked quite hard to engineer fair and serviceable plans for maintenance, resulting in the current offerings from TI for all university sites.

As well as working on these specific problems, we have had many meetings with Texas Instruments representatives wherein we have attempted to present the needs of the national community for short- and long-term AI workstation products, covering issues including the desirability of specialized hardware, address space, programming environment versus execution speed, and the ability to utilize the AI workstation's power for routine tasks.

Of course, there is also a large number of day-to-day activities needed to keep the computing environment pleasant, including resource management (e.g., disk space allocation, printer management), assistance with file backup and magnetic tape usage, and introducing new users to the system. We have produced documents targeted at complete novice users, users of InterLisp-D machines, and users of Symbolics machines in order to facilitate user education. These documents have been used as examples at various places in the national community.

For the coming year we plan to continue development and maintenance of the software tools, perhaps adding tools such as a DARPA Internet Domain Resolver, text processing facilities such as TeX, LaTeX, and document previewing tools, as well as aiding the growth of the users' group.

9.5 - Symbolics

Symbolics

Our work with Symbolics equipment has been slowed pending resolution of long-standing maintenance issues. As has been stated previously, in order for workstations to be competitive with time-shared mainframe computing resources, they must not only have a low purchase price, but must be cost-effective to maintain. This goal is normally achieved due to the economies of scale associated with having a large number of identical parts in an installation, as well as amortizing the cost of software development over many machines. We have come to reasonable agreements with all of the workstation vendors except for Symbolics. The high costs of service, the exceptionally high price of mail-in board repair, and the lack of a reasonable self-service alternative has left us unable to justify continued support of these machines unless a workable agreement can be reached. We have negotiated a tentative hardware maintenance contract, involving parts from Symbolics, and in-house labor, and are in negotiations for software maintenance. If we can reach consensus, we will be able to increase support of Symbolics machines once again.

While there have been no appreciable system development activities with the Symbolics machines, they have been maintained in good working order, with up-to-date software. We have not, however, moved from Release 6.1 to Genera 7.0 as the user community felt that the disadvantages of the transition overwhelmed the advantages, since the new software was quite incompatible with existing code, was slower, and seemed to introduce many new problems. We will re-evaluate Genera 7.1 when it is released. KEE and Fortran have been kept current, and patches in bulletins from Symbolics have been applied.

9.6 - SUN

We are just now bringing up several SUN workstations configured for Lisp research work. Several SUMEX projects have been able to experiment with SUN workstations through collaborations with other groups and the Lisp programming and debugging environments of these machines is still rather primitive as compared to the InterLisp and ZetaLisp machine environments. Also, SUN's need to be configured with relatively large memories to accommodate Lisp systems (because of limited garbage collection facilities currently) and this has required using third part memory. More standard configurations should be available from vendors shortly and we expect to have additional information to report next year.

10 - Workstation Standards and Access

10.1 - Computing Environment Standards

In a heterogeneous computing environment, such as AI research inevitably involves, the issue of cross-system compatibility is a central one. Users of various machines want to be able to share software, as well as be able to use various machines with a minimum of overhead in learning the operating procedures and programming languages of new systems. Thus, it is crucial to specify and propagate powerful, flexible standards for various aspects of the computing environment so that it is possible to transfer both skills and information among machines.

In order to improve the inter-machine compatibility of our software, we have been encouraging all users to use the CommonLisp programming language [16], as well as pressing vendors to provide more complete and efficient implementations of this language. We have already served as beta test sites for Xerox, Texas Instruments, and Lucid CommonLisp implementations.

The CommonLisp language, however, is only a subset of the software needed for our research. Research projects need higher-level powerful facilities, such as an object-oriented programming system and sophisticated error handling. Therefore we have been supporting and following the development of the CommonLisp Object System (CLOS) via membership in the electronic discussion group, technical contributions, and porting of Portable Common Loops (PCL), a predecessor of CLOS, to the TI Explorer. We are now encouraging vendors to produce efficient implementations of the system, and users to familiarize themselves with it. We are also encouraging vendors to adopt the proposed CommonLisp error system.

Other features of the computing environment also need to be standardized to be useful on more than one machine at a time. Another of the most important of these is the keyboard and display interface, often referred to as the "window system". See the virtual graphics section (page 34) for further discussion of window systems.

There are also many other areas which could benefit greatly from standardization, including document page description languages, text and graphics representations, and more networking protocols. However, it is important that standards not be entered into hastily, as an insufficient standard can often be worse than no standard at all. We intend to continue working to develop standards for these and other computing needs as the understanding of the issues involved matures.

10.2 - Protocol Standards

In addition to various portions of the AI research computing environment, the most highly visible area of standardization has been inter-machine communication, or networking. Underlying all network I/O must be a network protocol for packet transfer between cooperating hosts. At SUMEX we have had long term experience with several such protocols; PUP/BSP, PUP/EFTP, IP/TCP, IP/TFTP, IP/UDP, IP/SMTP, and NS/SPP are those most commonly used on SUNet. PUP/BSP and IP/TCP have been used to implement both FTP and TELNET, PUP/EFTP is an "Easy File Transfer Protocol" on top of PUP used for boot like services. IP/TFTP is a "Trivial File Transfer Protocol" which uses IP/UDP datagrams. IP/SMTP is the "Simple Mail Transfer Protocol" for sending mail, and runs on top of IP/TCP. NS/SPP is a "Sequenced Packet Protocol" similar to PUP/BSP and is used for FTP and TELNET. In the past we have elected to write servers for each new protocol in order to accommodate both vendor hardware and systems software requirements. This was necessary because no one protocol has been supported on all such systems.

With others in the computer science research community, we have pressed vendors to supply implementations of the DARPA standard TCP/IP communications protocols. We are pleased that the IP protocol family is now supported on all hardware and operating system configurations currently at SUMEX. And we expect to have IP support on any new systems we purchase in the future. Similarly, IP is supported on all of our UNIX based file servers, and the SUNet gateways route all IP datagrams. There has been a great deal of deliberate effort at Stanford and SUMEX to enforce IP as a standard protocol for new software development. This was motivated by its broad acceptance and the growing number implementations throughout the networking and vendor communities. This does not imply that we will abandon the other protocols but rather, since we are seeking to have *uniformity across all vendors* with the proposed Stanford distributed environment, we are choosing to limit new implementations to the IP protocol family. We are also currently working to provide improved support for TCP/IP in our Terminal Interface Processors (TIP's), having already implemented TCP/IP routing service.

As an example of the power of using uniform communication protocols, we set up a Xerox 1186 workstation for use by Dr. Shortliffe during his sabbatical in Philadelphia at the University of Pennsylvania. This university has a different network environment than Stanford's, although it is probably more typical of common Ethernet installations. The Pennsylvania network provides only Class-B IP/TCP services for VAX-based VMS/Unix systems over "thin" Ethernet. The 1186 was the only piece of Xerox hardware on the network so the disk was pre-loaded at Stanford. We successfully used the Pennsylvania VMS VAX's as file servers and time servers (after writing appropriate software to interface the workstation to the RFC868 time protocol). Using their Ethernet to ARPANET gateway, we were able to connect to SUMEX-AIM directly from the Xerox workstation as well as access our print servers at Stanford. Unfortunately, hardware problems with the workstation later in the year prevented us from attempting any more complex experiments with distributed computing and remote hosts.

Such standardization has a price, however, in that observed network communications speeds are uniformly higher between equipment "tuned" to individual vendor protocols. For a discussion of network file access protocol benchmarks see page 40.

11 - Network Services

A highly important aspect of the SUMEX system is effective communication within our growing distributed computing environment and with remote users. In addition to the economic arguments for terminal access, networking offers other advantages for shared computing. These include improved inter-user communications, more effective software sharing, uniform user access to multiple machines and special purpose resources, convenient file transfers, more effective backup, and co-processing between remote machines. Networks are crucial for maintaining the collaborative scientific and software contacts within the SUMEX-AIM community.

11.1 - Remote Networks

11.1.1 - Commercial Network Link

At the beginning of this grant year, SUMEX had just begun switching public data networks (PDN) vendors (from TYMNET to UNINET) in an attempt to improve service for our users. As the result of a corporate merger, our connection to UNINET became a connection to TELENET.

11.1.2 - X.25/Ethernet Link

SUMEX and Stanford's heavy use of Ethernet has prompted interest in a suitable connection between our Ethernet system and the Public Data Networks (specifically TELENET). Commercial groups provide a wide variety of equipment connecting these X.25 networks to Ethernets, but lack of standards for terminology make it difficult to determine their function.

Because our interest involves connection to other X.25 based hosts and Packet Assemblers/Disassemblers (PADs), sometimes called Terminal Interface Processors, as opposed to connecting two Ethernets via an X.25 net, we need a device that provides protocol translation. One alternative we are considering is to use a SUN processor for this task. This processor will have its normal TCP/IP Ethernet capability supplemented by an X.25 package provided by SUN. Such a package will provide SUMEX with both an inbound and outbound capability relative to TELENET. Users on SUMEX will be able to access the large variety of hosts and services on the PDNs (such as NLM and Dialog) in a simple and reliable manner. Though the high level protocols for file transfer and mail exchange are developing slowly in the X.25 environment, some progress is being made, so a general purpose interface to these networks is an important asset.

11.1.3 - ARPANET Link

We also continue our extremely advantageous connection to the Department of Defense's ARPANET, managed by the Defense Communications Agency (DCA). This connection has been possible because of the long-standing basic research effort in AI within the Knowledge Systems Laboratory that is funded by DARPA. ARPANET is the primary link between SUMEX and other university and AIM machine resources, including the large AI computer science community supported by DARPA. We are also attempting to establish a link to the DARPA wideband satellite network to facilitate the rapid transfer of large amounts of data such as are involved with projects like our Concurrent Symbolic Computing Architectures project.

As a member of the ARPAnet group, we have an obligation to help with certain network operations tasks. For instance, we participated in the upgrading (to 56 Kbs) of the connection which Advanced Decision Systems, Inc. (Mt. View, CA) has to our IMP. We also have a minor role in certain mail routing functions for the ARPANET community.

As part of an overall increase in ARPANET capacity a third 56 Kbs trunk line is being added to our IMP by the Defense Communications Agency (DCA).

11.2 - Microcomputer Networks

We connected our Apple Macintosh computers in 2 buildings with *Appletalk* and *Phonenet* network products. More significantly, we integrated them with the rest of our equipment by connecting the microcomputer networks to the campus Ethernet networks using *Kinetics FastPath* gateways, a commercial spinoff resulting from the SUMEX work on the SEAGATE gateway.

Software written at Columbia University, Stanford, and elsewhere, makes it possible for a Macintosh to share a VAX file server with the Lisp machines and to access hosts on the ARPA internet as a first-class workstation.

Usage of a centralized VAX file server makes nightly backup and data sharing automatic. This mode of usage is a great improvement over the isolated stand-alone machines that most people think of when they think of microcomputers.

11.3 - Local Area Networks

For many years now, we have been developing our local area networking systems to enhance the facilities available to researchers. Much of this work has centered on the effective integration of distributed computing resources in the form of mainframes, workstations, and servers. Network gateways and terminal interface processors (TIP's) were developed and extended to link our environment together. We are developing gateways to interface other equipment as needed too. A diagram of our local area network system is shown in Figure 8 and the following summarizes our LAN-related development work.

11.3.1 - Ethernet Gateways

In our heterogeneous network environment, in order to provide workstation access to file servers, mail servers, and other computers within the network, it is necessary to able to route multiple networking protocols through the network gateways. As reported last year, the SUMEX gateways support PUP, Xerox NS, Symbolics/Texas-Instrument CHAOSNET, and the IP/TCP protocols. This support not only provides the routers necessary to move such packets among the subnetworks, but also other miscellaneous services such as time, name/address lookup, host statistics, boot strap support, address resolution, and routing table broadcast and query information.

This year, with the acquisition of a SMI SUN 3/180 file server and three SUN 3/75 workstations, it was necessary to add special boot-protocol support for SMI's Net Disk and NFS protocols to allow the SUN workstations to boot their Unix kernel, and runnable programs while residing on a network that is distinct from the one on which the file server resides. SUN's convention is that each subnet must have its own file server that can provide boot support. But this is too expensive for complex network environments such as ours. Given this broadened capability, we can now place our "diskless" SUN workstations anywhere within the KSL network topology, rather than on the same network that the server resides.

Also, to improve the throughput of our highly loaded gateways, portions of Ethernet interface drivers and protocol routers were rewritten. The drivers now look for the arrival of additional packets while processing those packets that initiated the interrupt. Now, each router can process up to six packets before relinquishing control to the gateways process scheduler. Previously, each router would process only one such packet per call. These two changes more than doubled the maximum observed packets per second, as well as the maximum throughput bandwidth which is now about 2.5 megabits per second, and minimized the dropping of back-to-back packets by the Ethernet interface itself.

Over the past year our network topology grew in complexity and extent so that we now have redundant routes to several networks within the KSL and Stanford LAN. Within this more complex environment, the old routing table management schemes broke down and had to be redesigned and changed to adequately deal with the network interactions that arose. In particular, we had to ensure that when a route to a particular network no longer was available because of electrical, hardware or software failure, that this information was propagated throughout the topology in a manner that maintained routing table equilibrium. We have solved this problem and our gateways now recover gracefully from these failures.

A second kind of failure occurs when a path between two networks fails but the gateways involved are not aware of this fact, and as a consequence continue to advertise routes using paths that are partitioned. We have had two examples of this over the past year caused by the failure of a repeater in one case and a transceiver in the other. When we detect such a situation, we can now remove the route from the gateway generating it using software, make the repair, and then replace the route, without perturbing the connectivity of our topology if there are redundant routes around the partition caused by hardware failure.

Finally, a minor change in the gateways' routing table update algorithm when multiple routes to a network are available has managed to balance the load between these alternative paths, and increase the throughput at the gateways involved. Such gateways are usually focal points for high traffic volume, and the change was immediately noted by staff members sensitive to network throughput. The old version of the routing update protocol would hold onto a route even if alternative paths of equal cost were available. The new version will always update a route if a path of equal cost arises. When n redundant paths are available, the route changes approximately every 30/n seconds.

These services are still unique within the SUMEX-AIM portion of the Stanford University network, and give our researchers a networking environment that is flexible, of high bandwidth, and extremely dependable.

11.3.2 - Terminal Interface Processors

With the advent of reliable multiple speed (300, 1200 and 2400 baud) modems, we placed ten such devices on our TIPs for dial-in access, and added autobaud recognition to the TIP software. 2400 baud dial-in connections have shown themselves to be highly responsive in such a configuration, and have the advantage when placed on the TIP of giving the user access to any host on the Stanford local area network. Autobaud recognition has also been added to the directly attached tty ports to simplify user/TIP interaction. If a user changes his terminal's baud rate, the TIP will still be responsive, but at a different speed. Previously, such a line's baud rate was fixed, and this often led to a great deal of user frustration.

Also, the experimental NTT ELIS Lisp Machines used in the KSL currently do not have Ethernet connections. To accommodate remote access to these systems, they were attached to TIP ports so that a user could connect to the TIP from the Ethernet, and then transparently connect to the ELIS machines via a TIP command. Once this connection is established, the user appears to have a terminal directly attached to the ELIS itself. Currently, there are six such ELIS ports in use on one of our TIPs. Incidentally, the same code is currently being generalized for use as a dial-out module.

12 - Printing Services

Laser printers have become essential components of the work environment of the SUMEX-AIM community with applications ranging from scientific publications to hardcopy graphics output for ONCOCIN chemotherapy protocol patient charts. We have done much systems work to integrate laser printers into the SUMEX network environment so they would be routinely accessible from hosts and workstations alike. This software has been widely shared with other user groups in the AIM community and beyond.

SUMEX operates 7 medium-speed (8-20 pages per minute) Imagen laser printers, 2 low-speed (~3 ppm) Xerox laser printers, and 1 low-speed (~3 ppm) Apple laser printer. Each of the Imagen printers possesses an emulator for a line printer, a daisy wheel printer, a Tektronix plotter, and a typesetter (using the *Impress* language). The last 3 printers render the special-purpose *Press*, *Interpress*, and *Postscript* typesetter languages. In total, the laser printers printed about half a million pages of output during the year. Most of the printout was simple text, followed in quantity by formatted text in *Impress* format, Impress-format drawings, and screen dumps. Lastly, about 2000 pages each of Postscript-format drawings and formatted text were printed on the Apple Laser Writer. Although the *Postscript* language is probably the most popular typesetting language among commercial applications developers at the present time (and one which we support with the Laser Writer), the overwhelming preponderance of readily-renderable line printer and Impress jobs in our printing mix provides the basis for our decision to emphasize the relatively high-speed Imagen laser printers. Because of the increasing usage of Postscript among vendors, however, we have purchased an additional Apple Laser Writer for use in the Medical School Office Building.

In order to finally obtain families of fonts in common between our Press, Impress and Interpress printers, we used the TypeFounder software that we beta-tested for Xerox to extract font width information (for use by our workstations) from our existing Interpress printer fonts (a 12 page per minute, 300 dpi printer based on the Xerox 8000 processor) and also made new fonts using character splines from an earlier Xerox grant program. Having an overlap in fonts among the printers helps to relieve the problems inherent in trying to print the same complex document on different printer technologies. Some of the font additions required software patches for the Interpress driver software on the workstations. The Interpress driver was further modified to provide rotated fonts in order to print our specialized medical forms.

13 - General User Software

We have continued to assemble (develop where necessary) and maintain a broad range of user support software. These include such tools as language systems, statistics packages, vendor-supplied programs, text editors, text search programs, file space management programs, graphics support, a batch program execution monitor, text formatting and justification assistance, magnetic tape conversion aids, and user information/help assistance programs.

A particularly important area of user software for our community effort is a set of tools for inter-user communications. We have built up a group of programs to facilitate many aspects of communications including interpersonal electronic mail, a "bulletin board" system for various special interest groups to bridge the gap between private mail and formal system documents, and tools for terminal connections and file transfers between SUMEX and various external hosts. Examples of work on these sorts of programs have already been mentioned in earlier sections, particularly as they relate to extensions for a distributed computing environment.

At SUMEX-AIM we are committed to 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 availability of fast and convenient communication facilities coupling communities of computer facilities has made possible effective intergroup cooperation and decentralized maintenance of software packages. The many operating system and system software interest groups (e.g., TOPS-20, UNIX, D-Machines, network protocols, etc.) that have grown up by means of the ARPANET have been a good model for this kind of exchange. The other major advantage is that as a by-product of the constant communication about particular software, personal connections between staff members of the various sites develop. These connections serve to pass general information about software tools and to encourage the exchange of ideas among the sites and even vendors as appropriate to our research mission. We continue to import 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 assisted groups that have interacted with SUMEX user projects get access to software available in our community (for more details, see the section on Dissemination on page 103).

III.A.3.5. Relevant Core Research Publications

The following is a list of new publications and reports that have come out of our core research and development efforts over the past year:

KSL 85-57
(Journal Memo) E. Horvitz and D. Heckerman; **The Inconsistent Use of Measures of Certainty in Artificial Intelligence Research,** August 1985. To appear in: *Uncertainty in Artificial Intelligence* **15 pages**

KSL 85-58
(Journal Memo) C.D. Lane, M.E. Frisse, L.M. Fagan, and E.H. Shortliffe; **Object-Oriented Graphics in Medical Interface Design,** December 1985. To appear in: *AAMSI-86* **5 pages**

KSL 85-59
(Working Paper) Allan Terry; **Using Explicit Strategic Knowledge to Control Expert Systems,** December 1985. Submitted for publication in: *Artificial Intelligence* **51 pages**

KSL 85-60
(Working Paper) Jean-Luc Bonnetain; **FLOWER: A First Cut at Designing a Budget Proposal,** September 1985. **28 pages**

KSL 86-18
STAN-CS-86-1123. H. Penny Nii; **Blackboard Systems,** June 1986. To appear in: *AI Magazine Vols. 7-2 and 7-3.* **86 pages**

KSL 86-24
(Journal Memo) M.A. Musen, L.M. Fagan, D.M. Combs, and E.H. Shortliffe; **Using a Domain Model to Drive An Interactive Knowledge Editing Tool,** September 1986. To appear in: *Proceedings of AAAI Workshop on Knowledge Acquisition, 1986* **12 pages**

KSL 86-25
(Journal Memo) E.J. Horvitz, D.E. Heckerman, and C.P. Langlotz; **A Framework for Comparing Alternative Formalisms for Plausible Reasoning,** May 1986. **5 pages**

KSL 86-28
(Working Paper) James Brinkley, Craig Cornelius, Russ Altman, Barbara Hayes-Roth, Olivier Lichtarge, Bruce Duncan, Bruce Buchanan, Oleg Jardetzky; **Application of Constraint Satisfaction Techniques to the Determination of Protein Tertiary Structure,** March 1986. **14 pages**

KSL 86-29
(Working Paper) Matthew L. Ginsberg; **Multi-valued logics,** April 1986. To appear in: *AAAI - 86* **13 pages**

KSL 86-33
(Journal Memo) David E. Heckerman and Eric J. Horvitz; **The Myth of Modularity in Rule-Based Systems,** May 1986. **7 pages**

KSL 86-36
STAN-CS-87-1148. Bruce A. Delagi, Nakul Saraiya, Sayuri Nishimura, and Greg Byrd; **An Instrumented Architectural Simulation System,** January 1987. **21 pages**

KSL 86-37
(Working Paper) Matthew L. Ginsberg; **Possible Worlds Planning,** April 1986. Submitted for publication to: *1986 Planning Workshop* **13 pages**

KSL 86-38
STAN-CS-87-1147. Barbara Hayes-Roth, M. Vaughan Johnson Jr., Alan Garvey, and Michael Hewett; **A Modular and Layered Environment for Reasoning about Action,** April 1987. To appear in: *The Journal of Artificial Intelligence in Engineering, Special Issue on Blackboard Systems, October 1986.* **63 pages**

KSL 86-39
(Journal Memo) E.H. Shortliffe; **Artificial Intelligence in Management Decisions: ONCOCIN,** April 1986. To appear in: *Proceedings of a Conference on Medical Information Sciences, University of Texas Health Sciences Center at San Antonio, July 1985. Also in Frontiers of Medical Information Sciences, Praeger Publishing, 1986.* **14 pages**

KSL 86-40
(Journal Memo) Christopher Lane; **The Ozone Manual,** July 1986. **34 pages**

KSL 86-42
(Working Paper) Oleg Jardetzky, Andrew Lane, Jean-Francois Lefevre, Olivier Lichtarge, Barbara Hayes-Roth, Russ Altman, Bruce Buchanan; **A New Method for the Determination of Protein Structures in Solution from NMR,** May 1986. Submitted for publication in: *Proc. XXIII Congress Ampere, Rome, Italy, Sept. 1986* **6 pages**

KSL 86-43
(Journal Memo) Edward H. Shortliffe; **Update on Oncocin: A Chemotherapy Advisor for Clincal Oncology,** August 1986. Submitted for publication in: *Medical Informatics* **4 pages**

KSL 86-44
(Thesis) Stephen M. Downs; **A Program for Automated Summarization of On-Line Medical Records,** June 1986. **27 pages**

KSL 86-46
STAN-CS-86-1111. Paul Rosenbloom and John Laird; **Mapping Explanation-Based Generalization onto Soar,** June 1986. To appear in: *AAAI-86* **18 pages**

KSL 86-47
STAN-CS-86-1124. Daniel J. Scales; **Efficient Matching Algorithms for the SOAR/OPS5 Production System,** June 1986. **50 pages**

KSL 86-48
(Working Paper) William J. Clancey; **Review of Winograd and Flores' "Understanding Computers and Cognition: A New Foundation for Design",** July 1986. **13 pages**

KSL 86-49
(Journal Memo) M.A. Musen, D.M. Combs, J.D. Walton, E.H. Shortliffe, L.M. Fagan; **OPAL: Toward the Computer-Aided Design of Oncology Advice Systems,** July 1986. Submitted for publication to: *Proceedings of the Tenth Annual Symposium on Computer Applications in Medical Care.* **10 pages**

KSL 86-50
(Working Paper) Ross D. Shacter and David E. Heckerman; **A Backwards View for Assessment,** July 1986. **6 pages**

KSL 86-51
Barbara Hayes-Roth, Bruce Buchanan, Olivier Lichtarge, Michael Hewett, Russ Altman, James Brinkley, Craig Cornelius, Bruce Duncan, and Oleg Jardetzky; **PROTEAN: Deriving protein structure from constraints,** March 1986. To appear in: *Proceedings of AAAI 1986* **21 pages**

KSL 86-52
(Working Paper) Edward H. Shortliffe, M.D., Ph.D; **Medical Expert Systems: Knowledge Tools for Physicians,** September 1986. **24 pages**

KSL 86-53
(Working Paper) Edward H. Shortliffe, M.D., Ph.D; **Medical Expert Systems Research at Stanford University,** September 1986. **13 pages**

KSL 86-56
(Working Paper) Nakul P. Saraiya; **AIDE: A Distributed Environment for Design and Simulation** June 1986. **25 pages**

KSL 86-57
(Working Paper) Curtis P. Langlotz, Edward H. Shortliffe, and Lawrence M. Fagan; **A Methodology for Computer-Based Explanation of Decision Analysis,** November 1986. **21 pages**

KSL 86-58
William J. Clancey; **Intelligent Tutoring Systems: A Tutorial Survey,** September 1986. Submitted for publication in: *Collected papers of the International Professorship in Computer Science (Expert Systems) Universite de L'Etat, Belgium* 43 pages

KSL 86-60
(Working Paper) Alan Garvey, Michael Hewett, M. Vaughan Johnson, Robert Schulman, Barbara Hayes-Roth; **BB1 User Manual - Interlisp Version,** October 1986. **68 pages**

KSL 86-61
(Working Paper) Alan Garvey, Michael Hewett, M. Vaughan Johnson, Robert Schulman, Barbara Hayes-Roth; **BB1 User Manual - Common Lisp Version,** October 1986. **72 pages**

KSL 86-62
(Working Paper) David C. Wilkins; **On the Limits of Debugging via Differential Modeling,** October 1986. **15 pages**

KSL 86-63
(Working Paper) David C. Wilkins; **Knowledge Base Debugging Using Apprenticeship Learning Techniques,** October 1986. **15 pages**

KSL 86-64
(Working Paper) Donald E. Henager; **Window-Driven Object-Oriented Calculator,** March 1986. **56 pages**

KSL 86-65
Matthew L. Ginsberg, David E. Smith; **Reasoning About Action I: A Possible Worlds Approach,** May 1987. **25 pages**

KSL 86-66
Matthew L. Ginsberg, David E. Smith; **Reasoning About Action II: The Qualification Problem,** May 1987. **28 pages**

KSL 86-68
David E. Smith; **Controlling Backward Inference,** March 1987. **67 pages**

KSL 86-69
STAN-CS-86-1136. Harold Brown, Eric Schoen, and Bruce Delagi; **An Experiment in Knowledge-Base Signal Understanding Using Parallel Architectures,** October 1986. To appear in: *Parallel Computation and Computers for AI, J.S. Kowalik Editor, Kluwer Publishers.* 39 pages

KSL 86-70
STAN-CS-86-1140. John E. Laird, Allen Newell, and Paul S. Rosenbloom; **Soar: An Architecture for General Intelligence,** December 1986. To appear in: *Artificial Intelligence.* **66 pages**

KSL 86-74
(Thesis) Glenn Douglas Rennels; **A Computational Model of Reasoning from the Clinical Literature,** June 1986. **244 pages**

KSL 86-75
(Journal Memo) Eric J. Horvitz; **Toward a Science of Expert Systems,** March 1986. **8 pages**

KSL 86-76
M. Vaughan Johnson Jr. and Barbara Hayes-Roth; **Integrating Diverse Reasoning Methods in the BB1 Blackboard Control Architecture,** December 1986. **17 pages**

KSL 87-01
(Working Paper) David C. Wilkins, William J. Clancey, and Bruce G. Buchanan; **Knowledge Base Refinement Using Abstract Control Knowledge,** January 1987. **9 pages**

KSL 87-02
STAN-CS-87-1146. Gregory T. Byrd, Russell Nakano, and Bruce A. Delagi; **A Point-to-Point Multicast Communications Protocol,** January 1987. **30 pages**

KSL 87-03
Bruce G. Buchanan; **Artificial Intelligence As An Experimental Science,** January 1987. To appear in: *Synthese* **41 pages**

KSL 87-05
STAN-CS-87-1142. James F. Brinkley, Bruce G. Buchanan, Russ B. Altman, Bruce S. Duncan, Craig W. Cornelius; **A Heuristic Refinement Method for Spatial Constraint Satisfaction Problems,** January 1987. **15 pages**

KSL 87-06
(Journal Memo) Glenn D. Rennels; **A Computational Model of Reasoning from the Clinical Literature,** January 1987. To appear in: *SCAMC Proceedings, Washington D.C. 1986* **8 pages**

KSL 87-07
STAN-CS-87-1144. Gregory T. Byrd and Bruce A. Delagi; **Considerations for Multiprocessor Topologies,** January 1987. **6 pages**

KSL 87-08
Robert Schulman and Barbara Hayes-Roth; **ExAct: A Module for Explaining Actions,** January 1987. **15 pages**

KSL 87-09
(Working Paper) Peter D. Karp and Peter Friedland; **Coordinating the Use of Qualitative and Quantitative Knowledge in Declarative Device Modeling,** January 1987. **20 pages**

KSL 87-11
Alan Garvey, Craig Cornelius, and Barbara Hayes-Roth; **Computational Costs versus Benefits of Control Reasoning,** February 1987. **13 pages**

KSL 87-12
(Working Paper) William J. Clancey; **The Knowledge Engineer as Student: Metacognitive bases for asking good questions,** January 1987. To appear in *Learning Issues for Intelligent Tutoring Systems, Heinz Mandl and Alan Lesgold, editors. Springer-Verlag: New York* **30 pages**

KSL 87-16
(Journal Memo) Eric J. Horvitz; **Inference under Varying Resource Limitations,** February 1987. **16 pages**

KSL 87-18
Isabelle de Zegher-Geets, Andy Freeman, Mike Walker, Bob Blum, Gio Wiederhold; **Computer-Aided Summarization of a Time-Oriented Medical Database,** February 1987. **6 pages**

KSL 87-19
(Working Paper) Curt P. Langlotz and Edward H. Shortliffe; **The Relationship between Decision Theory and Default Reasoning,** February 1987. **16 pages**

KSL 87-20
(Working Paper) Michael G. Kahn; **Model-Based Interpretation of Time-Ordered Data,** March 1987. **18 pages**

KSL 87-21
(Working Paper) Gregory F. Cooper; **An Algorithm for Computing Probabilistic Propositions,** March 1987. **5 pages**

KSL 87-22
(Journal Memo) Homer L. Chin and Gregory F. Cooper; **Stochastic Simulation of Casual Bayesian Models,** March 1987. **11 pages**

KSL 87-23
(Working Paper) Thierry Barsalou and Gio Wiederhold; **Applying a Semantic Model to an Immunology Database,** March 1987. **18 pages**

KSL 87-24
(Journal Memo) Homer L. Chin and Gregory F. Cooper; **Knowledge-Based Patient Simulation,** March 1987. **11 pages**

KSL 87-25
(Journal Memo) Edward H. Shortliffe; **Computers in Support of Clinical Decision Making,** March 1987. **12 pages**

KSL 87-32
(Working Paper) William J. Clancey; **Diagnosis, Teaching, and Learning: An Overview of GUIDON2 Research,** April 1987. 12 Pages

KSL 87-34
(Working Paper) Russell Nakano; **Experiments with a Knowledge-Based System on a Multiprocessor: Preliminary AIRTRAC-LAMINA Qualitative Results.** June, 1987.

KSL-87-35
(Working Paper) Masafumi Minami; [Experiments with a Knowledge-Based System on a Multiprocessor: Preliminary AIRTRAC-LAMINA Quantitative Results.] June, 1987.

Other Outside Articles:

Hayes-Roth, B., Johnson, M.V., Garvey, A., and Hewett, M. Applications of BB1 to arrangement-assembly tasks. *Artificial Intelligence in Engineering*, October, 1986.

Hayes-Roth, B., Johnson, M.V., Garvey, A., and Hewett, M. The BB* environment. To appear in: R. Engelmore and A. Morgan (Eds.), *Blackboard Systems*. Addison-Wesley, London: 1987.

Garvey, A., Cornelius, C., and Hayes-Roth, B. Computational costs versus benefits of control reasoning. *Proceedings of the American Association for Artificial Intelligence: AAAI87*, 1987.

M. V. Johnson, and Hayes-Roth, B. Integrating diverse reasoning methods in BB1. *Proceedings of the American Association for Artificial Intelligence: AAAI87*, 1987.

Hayes-Roth, B. Blackboard systems. In Eckroth, D. (Ed.), *Encyclopedia of Artificial Intelligence*. New York: John Wiley & Sons, 1987.

Garvey, A., Hewett, M., Johnson, M.V., Schulman, R., and Hayes-Roth, B. *BB1 User's Manual*. Stanford University. 1987.

Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewett, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O. Elucidating protein structure from constraints in PROTEAN. To be reprinted in: R. Engelmore and A. Morgan (Eds.), *Blackboard Systems*. Addison-Wesley, London: 1987.

Duncan, B., Buchanan, B. G., Hayes-Roth, B., Lichtarge, O., Altman, R., Brinkley, J., Hewett, M., Cornelius, C., and Jardetzky, O. PROTEAN: A new method of deriving solution structures of proteins. *Bulletin of Magnetic Resonance*, 1987.

Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewett, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O. PROTEAN: Deriving protein structure from constraints. *Proceedings of the American Association for Artificial Intelligence: AAAI86*, 1986.

Jardetzky, O., Lane, A., Lefevre, J., Lichtarge, O., Hayes-Roth, B., Altman, R., and Buchanan, B. A new method for the determination of protein structures in solution from NMR. In *Proceedings of the XXIII Congress Ampere*, Rome, Italy, 1986.

Tommelein, I. D., Johnson, M. V., Hayes-Roth, B., and Levitt, R.E., SIGHTPLAN: A Blackboard Expert System for the Layout of Temporary Facilities on a Construction Site. *Proceedings of the IFIP WG5.2 Conference on Expert Systems in Computer-Aided Design*, Sydney, Australia, 1987.

Tommelein, I. D., Levitt, R. E., and Hayes-Roth, B., Using Expert Systems for the Layout of Temporary Facilities on Construction Sites." *Proceedings of the CIB W-65 Symposium on Organization and Management of Construction*, Birkshire, U.K, 1987.

III.A.3.6. Resource Equipment

The SUMEX-AIM core facility, started in March 1974, was built around a Digital Equipment Corporation (DEC) KI-10 computer and the TENEX operating system and continued through the 1970's with a mainframe focus for the resource. The interactive computing environment of this facility, with its AI program development tools and its network and interpersonal communication media, was unsurpassed in other machine environments. Biomedical scientists found SUMEX easy to use in exploring applications of developing artificial intelligence programs for their own work and in stimulating more effective scientific exchanges with colleagues across the country. Coupled through wide-reaching network facilities, these tools provided us access to a large computer science research community, including active artificial intelligence and system development research groups.

In the late 1970's and early 1980's, computer system research on early microprocessors and compact minicomputers suggested that large mainframe computers would not be essential or even the dominant source of computing power for AI research and AI program dissemination. Thus, we began to implement a strategy for computing resources marked by the integration of heterogeneous systems -- mainframes, Lisp workstations, and service systems (e.g., for file storage and printing) all linked together by local area networks. Over the years, we have configured the optimal resource computing environment around shared central machines coupled through a high-performance network to growing clusters of personal workstations.

The concept of the individual workstation, especially with the high-bandwidth graphics interface, proved ideal. Both program development tools and facilities for expert system user interactions were substantially improved over what is possible with a central time-shared system. The main shortcomings of early workstation systems were their limited processing speed and high cost. But in the few years since our first experimental systems, processing power has increased by more than a factor of 10 and the cost has decreased by a comparable factor.

Today the SUMEX resource is a complex, integrated facility comprised of machines, networks, and servers illustrated in Figures 4 - 8. A key role of the SUMEX-AIM resource is to continue to evaluate workstations as the technology is changing rapidly. This evaluation includes new hardware and software, 1) to provide superior development and execution platforms for AI research, and 2) to support the ancillary "office environment" (presently carried out on the DEC 2060, which is being phased out). Thus far no single workstation has materialized that provides all the services we would like to see in support of either or both of these missions. This means that for the foreseeable future, we will utilize a multiplicity of machines and software to address the needs of the projects.

Systems based on the Motorola 68020 chip (e.g., SUN Microsystems or Apple Macintosh II workstations), the Intel 80286 and 80387 chips (e.g., IBM PS/1-4 machines), and other newer architectures, such as reduced instruction set computer (RISC) chips, have Lisp benchmark data rivaling the performance of existing, specially microcoded Lisp machines (e.g., those from Xerox, Symbolics, and TI). But these Lisp machine vendors are producing substantially faster machines as well, using VLSI technology. It is still too early to predict how this "race" will ultimately turn out and software environments will play an equally important role to raw hardware speed in the decision. For now, the Lisp software environments on the "stock" machines are not nearly so extensively developed as on Lisp machines and conversely, the routine computing environments of Lisp machines (text processing, mail, spreadsheets, etc.) lag the tools available on stock UNIX machines.

In the past year we experimentally tried increasing usage of TI and Xerox Lisp machines (purchased as AI research platforms) for text editing and document formatting, but their functionality and speed do not approach that of the *TeX* and *Scribe* formatters when executed on the 2060. The Lisp machines do not yet provide the myriad tools of the 2060 (e.g. mail, database, spreadsheet, dictionary), but we and other groups are undertaking to rewrite mainframe tools to address the most pressing shortcomings. Another problem is that execution of office tools on these machines impacts their utility as research tools. Improvements in processor speed, memory size, display size, and window systems may address this problem in the near future. We look forward to the introduction and testing of the TI *Explorer II* and Xerox *Tamarind* with this in mind.

Some community members tried increasing their usage of Macintosh applications as a means of reducing dependence on the 2060, but except for some drawing tools, they were not up to the job, hampered by the small display and incapacity to cope with large or complex documents. We look forward to investigating the much more powerful Macintosh II as an office system and possibly Lisp delivery vehicle. Early indications suggest limited potential for Lisp development, but perhaps mass availability will encourage improvement in this area.

In the long term, we may hope to see an integration of both the Lisp machine and stock machine worlds. Despite the inadequacy of the present single-vendor offerings, the potential leverage of Lisp machine technology for office systems ancillary to research makes the pursuit of combining the two as attractive as ever, and we intend to take advantage of new hardware opportunities as they arise.

1 - Purchases This Past Year

The core resource hardware continues to be stable and the relatively small amount of SUMEX-AIM money for new purchases has been concentrated on experimental workstations and server equipment needed for distributed system development. These purchases are paced carefully with the developments of higher performing, more compact, and lower cost systems. The purchases this past year are summarized below. It should be noted that these purchases in many cases complement hardware acquired with non-NIH funding, including 3 SUN 3/75 workstations, a SUN 3/180 file server, and numerous laser printer upgrades.

1. SUN X-501B 75 Megabyte Disk Drives (3 each, for Lisp workstations)
2. Sun 6250 BPI Tape Drive (for file server backup)
3. Parity 24-Megabyte memory boards (3 each, for Lisp Workstations)
4. Apple Macintosh SE computer (for text processing and graphics)
5. Apple Mac II computer (40 Megabyte disk, 7 MB memory upgrade, and video card/monitor; for a Lisp workstation)
6. Imagen 3320-3 laser printer (for higher volume printing)
7. Ricoh 4120 laser printer (used, for spare parts)
8. Toshiba T1100 Plus Portable Computer (as a portable travel computer)
9. Ethernet (10MB bits) Multibus Interface Boards (4 each, for network expansions)
10. U.S. Robotics 9600 baud modems (2 each, for higher speed serial line connections)

1.1 - Workstation Hardware

Using non-DRR funding, the KSL has taken delivery of 20 new Xerox 1186 LISP workstations, and has upgraded 4 Xerox 1108 machines to 1109 (Dandetigers) with memory expansion and floating point support. The machines are used by many projects in the KSL, including the GUIDON and NEOMYCIN efforts, BB1, PROTEAN, and Financial Resources Management (FRM). These machines increase our research capabilities and complement the Texas Instrument Explorers and Symbolics 36XX facilities of the KSL.

Our Xerox workstations proved to be very reliable again this year and justified our strategy of saving money by not purchasing service contracts. Also to save money, we arranged with third parties to repair/replace some components that did fail. (Exception: We purchased a third-party service contract on one Xerox 1132 disk drive since the particular device has failed more than once.)

The basic components of the three Sun 3/75 workstations were purchased with DARPA funding for evaluation as AI development engines and/or office systems. Although Sun recommends these machines as general purpose workstations, experience indicated that memory and disk upgrades to the basic systems are necessary to consider their use as Lisp engines. These upgrades are on-order and evaluation is still in the early stages.

1.2 - File Server Hardware

Because our Lisp workstations have only limited local file space, the development of effective shared file servers is essential to our resource operation. SUMEX now has three UNIX-based file servers. Two of them, as reported in the past, use VAX/750's as the processors: the SAFE has four, 470 Megabyte, Fujitsu Eagle disk drives and the ARDVAX has one such disk drive. The SAFE also is equipped with a 300 megabyte CDC, removable media, disk drive and a 800/1600 BPI Kennedy tape drive. The CDC unit is used for incremental backup dumps and the tape drive is used for both incremental and full backup dumps. A procedure has been established whereby the ARDVAX is able to use this equipment for its incremental and full dumps over the network. The configurations of these systems are shown in Figure 7.

With DARPA funding this past year we bought a system called the KNIFE, a file server based on a SUN 3/180 processor. It is equipped with two of the 470 Megabyte Fujitsu disk drives and a cartridge tape drive (see Figure 5). We are in the process of adding a Fujitsu 1600/6250 BPI tape drive for backup dumping. Being relatively new, the performance of this equipment in an operational environment has not yet been thoroughly checked out at SUMEX.

The Xerox XNS Ethernet-based file server (donated by Xerox in 1985) has increased in capacity and usage in the past year. This server is based on the Xerox 8000 processor (identical hardware to the Xerox 1108 Lisp workstation but running more conventional microcode) and the Century Data Systems T-305 removable media disk drive. With the addition of two additional disk drives (also donated), the total potential storage capacity of the server has increased to approximately 900 MB (of which 600 MB is currently available from the network).

The user base for this server has grown to over sixty regular, registered users and numerous infrequent guest and project users. This server is the primary system software resource for over fifty Lisp workstations. In the past year, the server software has been upgraded twice, the most recent upgrade introduced random access to the content of files which, when interfaced to Interlisp's paged file mechanisms, should improve both the flexibility and effective speed of the server.

Though optical disks have been slow in realizing their earlier-announced potential, suitably packaged products are now appearing in the marketplace. It is possible that this technology used in place of (or in conjunction with) conventional magnetic tapes might provide an excellent medium for implementing a responsive offline storage system for data. It is fair to expect that even a small laboratory could have reasonable access to hundreds of gigabytes of storage.

1.3 - Printer Hardware

Over the past year, we purchased 2 new Imagen 12/300's, upgraded an 8/300 to a 12/300, and converted an old Hewlett-Packard 2688A to a 12/300 laser printer for the SUMEX-AIM community. These enhancements were funded by DARPA. The move to 12/300's was motivated primarily by the ruggedness of the Ricoh LP-4120 print engine used in those printers. Whereas the Canon LBP-CX print engine used in the 8/300 has an expected lifetime of 70,000 pages, the Ricoh LP-4120 has an expected lifetime of 700,000 pages. Other beneficial side-effects of the upgrade were: (1) higher print rate (12 pages-per-minute), (2) bigger paper tray (half a ream), (3) blacker and more solid print, (4) crisper print, and (5) cheaper supplies (half the price per page compared to the 8/300).

We have also acquired an Apple Laser Writer which interprets the PostScript page description language. Within a few months of its introduction, the Apple Laser Writer has become the most common laser printer on campus and around the world. Economies of scale have made it possible for us to acquire this printer for under $4000. SUMEX AppleNet/Ethernet expertise will make it possible for us to attach the Laser Writer to the high-bandwidth campus internet and operate the printer at the high-end of its 8 page-per-minute capacity. (The vast majority of laboratory-owned Laser Writers in the U.S. are driven over a low-bandwidth RS-232 line yielding only 3 pages-per-minute throughput and typically greater latency.) The PostScript page description language is already the standard of choice at university and DARPA sites (judging by traffic on the Laser-Lovers discussion group). It is generally agreed upon in these communities that PostScript is among the easiest-to-generate and most expressive of the page description languages in use today and reconciles these traits much more effectively than other languages do.

At present, most of our printers image at 300 dots per inch (dpi) and our finest printer is the aging Xerox Alto-Raven which images at 384 dpi. To exploit the special capabilities of much higher quality, camera-ready printers and to take advantage of the economical Apple Laser Writer, we have begun an Interlisp implementation of an "image stream" driver for PostScript. UNILOGIC has already added Postscript support to Scribe and Adobe has implemented Postscript support for TeX.

1.4 - Network Hardware

As we evolved a more complex network topology and decided to compartmentalize the overall Stanford internet to avoid electrical interactions during development and to facilitate different administrative conventions for the use of the various networks, we developed gateways to couple subnetworks together using Motorola MC-68000 systems. Given the heterogeneity of our environment, these gateways continually need to provide additional services to support the influx of new workstations. To accommodate current and anticipated gateway software growth, we have increased the memory capacity of the MC-68000 cpu board from 256 kilobytes to 1 megabyte.

We also developed a MC-68000 terminal interface processor (TIP) to provide terminal access to network hosts and facilities. It is basically a machine that has a number of terminal lines and a network interface and software to manage the establishment of connections for each line and the flow of characters between the terminal and host. In the past, 32 lines per TIP was sufficient, but our transition plan for moving users off the 2060 includes moving both the dial-in and dial-out functionality of the 2060 to TIPs, and this year we upgraded one of our TIPs to support 10 such ports. Thus, the 32 line upper bound is no longer feasible, and there is now the need to configure TIPs with at least 48, and perhaps 64 lines. As with the gateways, we have quadrupled the memory size of the TIPs' MC-68000 cpu board to 1 megabyte. This will adequately handle any future expansion of these servers. We have also improved the Dial IN/OUT service for both the 2060 and Tips for faster operation (2400 baud service maximum).

SUMEX-AIM is continuing its efforts in improving the networking environment for faster and more unified data communications. In this report period, several reconfigurations towards this endeavor have been completed. The SUMEX-AIM facility has been relocated to a new building. This move necessitated the relocation of all offices as well as all associated computer equipment. A network in the new building had to be designed and implemented and coupled into the old one which connects with the remaining KSL groups as well as the Stanford campus proper. This modification gave us the opportunity to upgrade several portions of the network in a manner that will provide redundancy as well as future expansion capabilities to the Medical Center and all other planned adjacent buildings. The new facility was wired to provide every sitting space with a flexible network connect capability similar to a telephone type connection. The entire scheme was successfully implemented with very little downtime. After almost a year in operation this scheme seems to be very reliable.

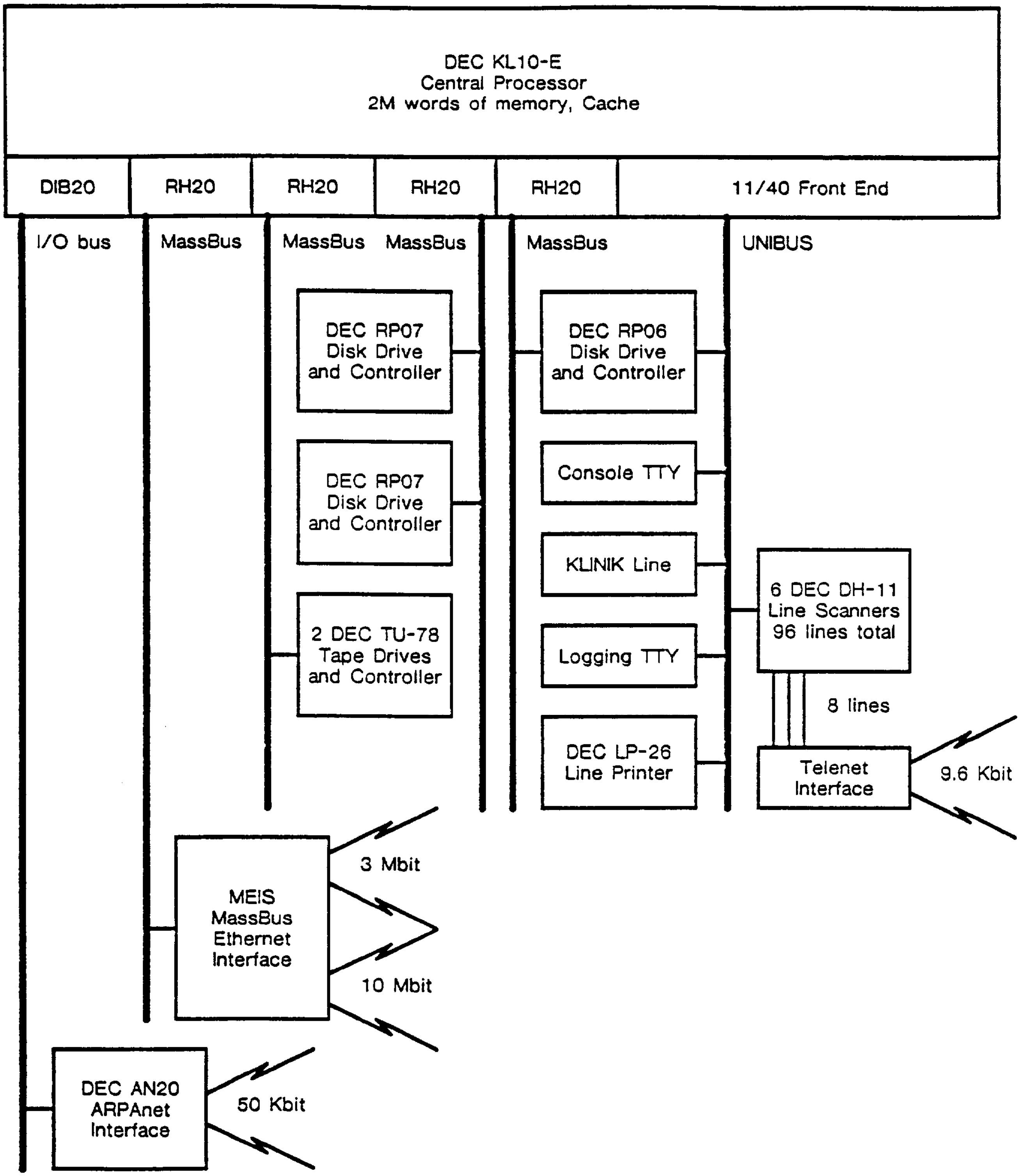


Figure 4: SUMEX-AIM DEC 2060 Configuration

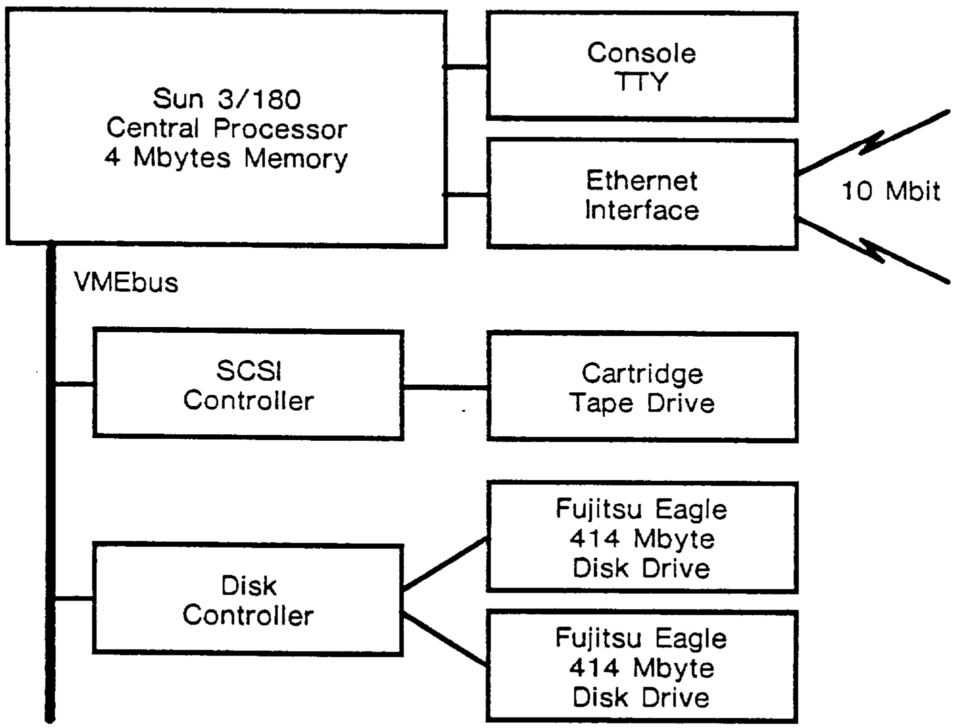


Figure 5: SUMEX-AIM Sun File Server Configuration

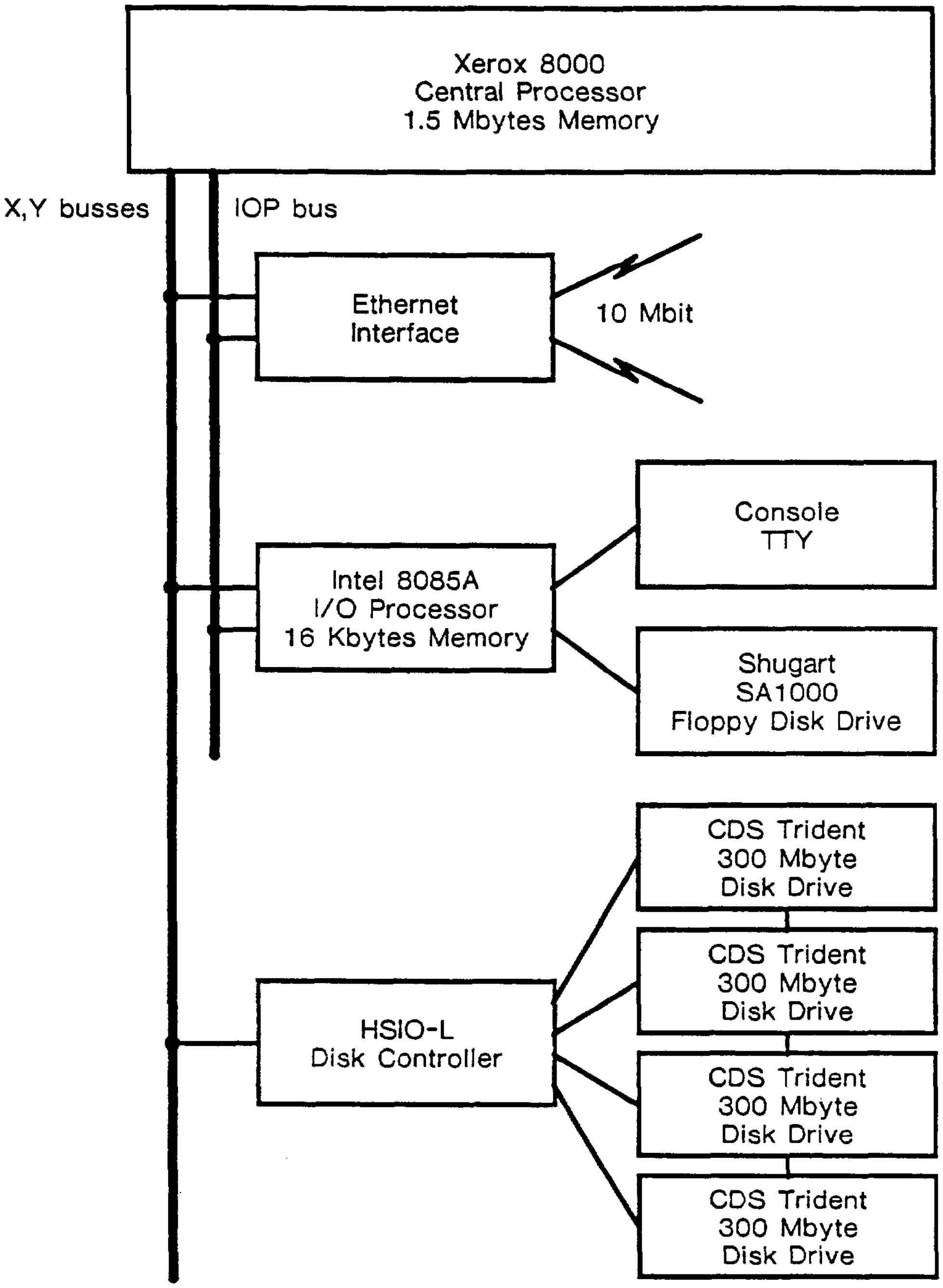


Figure 6: SUMEX-AIM Xerox File Server Configuration

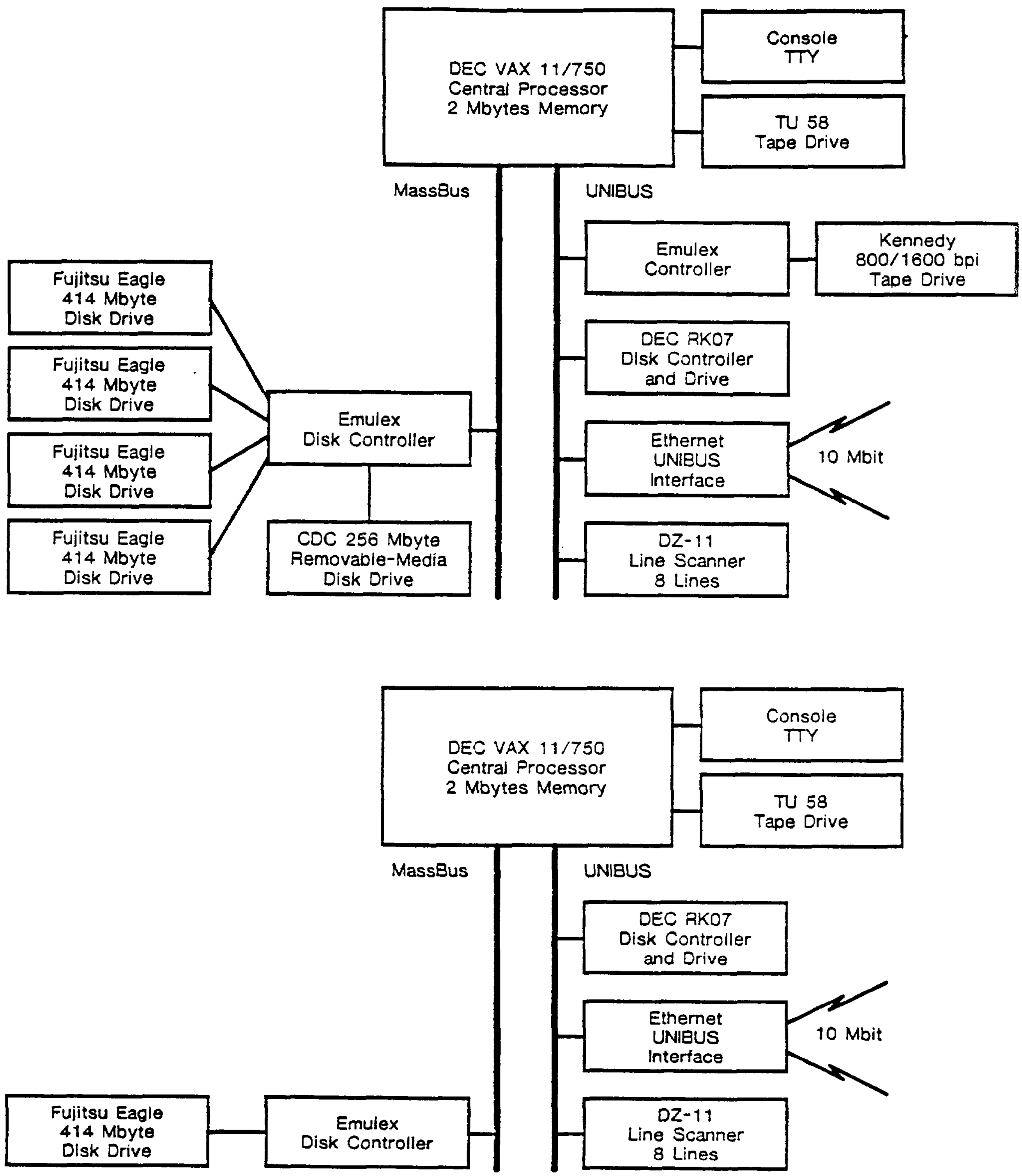


Figure 7: SUMEX-AIM VAX File Server Configuration

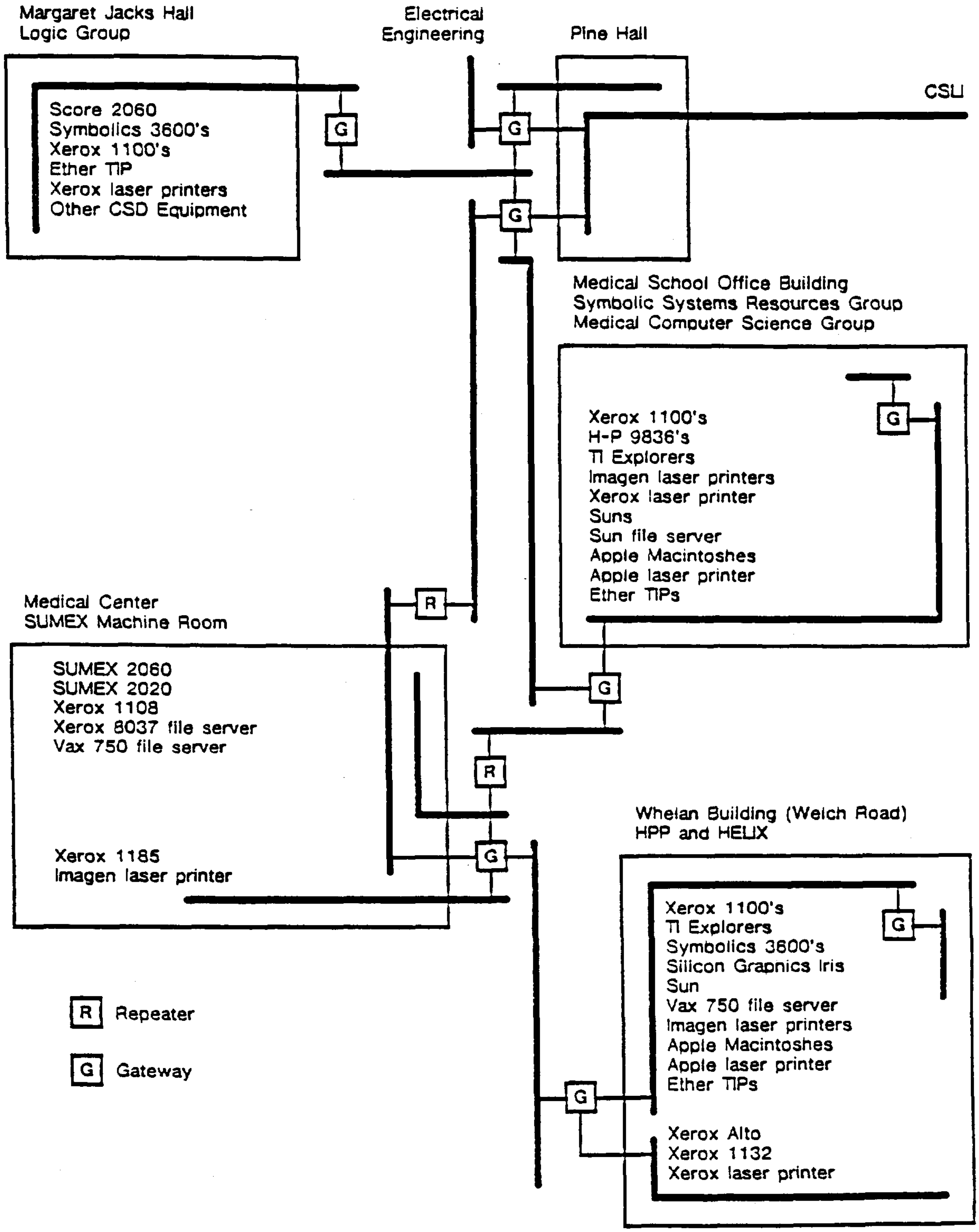


Figure 8: SUMEX-AIM EtherNet Configuration

III.A.3.7. Training Activities

The SUMEX resource exists to facilitate biomedical artificial intelligence applications. This user orientation on the part of the facility and staff has been a unique feature of our resource and is responsible in large part for our success in community building. The resource staff has spent significant effort in assisting users gain access to the central resource system and use it effectively as well as in assisting AIM projects in designing their own local computing resources based on SUMEX experience. We have also spent substantial effort to develop, maintain, and facilitate access to documentation and interactive help facilities. The HELP and Bulletin Board subsystems have been important in this effort to help users get familiar with the computing environment.

We have regularly accepted a number of scientific visitors for periods of several months to a year, to work with us to learn the techniques of expert system definition and building and to collaborate with us on specific projects. Our ability to accommodate such visitors is severely limited by space, computing, and manpower resources to support such visitors within the demands of our on-going research.

Finally, the training of graduate students is an essential part of the research and educational activities of the KSL. Based largely on the SUMEX-AIM community environment, we have initiated two unique, special academic degree programs at Stanford, the Medical Information Science program and the Masters of Science in AI, to increase the number of students we produce for research and industry. A number of students are pursuing interdisciplinary programs and come from the Departments of Engineering, Mathematics, Education, and Medicine.

The *Medical Information Sciences (MIS)* program continues to be one of the most obvious signs of the local academic impact of the SUMEX-AIM resource. The MIS program received recent University approval (in October 1982) as an innovative training program that offers MS and PhD degrees to individuals with a career commitment to applying computers and decision sciences in the field of medicine. In Spring 1987, a University-appointed review group unanimously recommended that the degree program be continued for another five years. The MIS training program is based in the School of Medicine, directed by Dr. Shortliffe, co-directed by Dr. Fagan, and overseen by a group of six University faculty that includes two faculty from the Knowledge Systems Laboratory (Profs. Shortliffe and Buchanan). It was Stanford's active on-going research in medical computer science, plus a world-wide reputation for the excellence and rigor of those research efforts, that persuaded the University that the field warranted a new academic degree program in the area. A group of faculty from the medical school and the computer science department argued that research in medical computing has historically been constrained by a lack of talented individuals who have a solid footing in both the medical and computer science fields. The specialized curriculum offered by the new program is intended to overcome the limitations of previous training options. It focuses on the development of a new generation of researchers with a commitment to developing new knowledge about optimal methods for developing practical computer-based solutions to biomedical needs.

The program accepted its first class of four trainees in the summer of 1983 and has now reached its steady-state size of approximately twenty graduate students. We do not wish to provide too narrow a definition of what kinds of prior training are pertinent because of the interdisciplinary nature of the field. The program has accordingly encouraged applications from any of the following:

- medical students who wish to combine MD training with formal degree work and research experience in MIS;
- physicians who wish to obtain formal MIS training after their MD or their

residency, perhaps in conjunction with a clinical fellowship at Stanford Medical Center;

- recent BA or BS graduates who have decided on a career applying computer science in the medical world;
- current Stanford undergraduates who wish to extend their Stanford training an extra year in order to obtain a "co-terminus" MS in the MIS program;
- recent PhD graduates who wish post-doctoral training, perhaps with the formal MS credential, to complement their primary field of training.

In addition, a special one-year MS program is available for established academic medical researchers who may wish to augment their computing and statistical skills during a sabbatical break. As of Spring 1987, half our trainees have previously received MD degrees and another quarter are medical students enrolled in joint degree programs. One-third are candidates for the MS degree, while the rest are doctoral students. The program has three graduates to date, with several more expecting to complete degrees before the end of 1987.

Except for the special one-year MS mentioned above, all students spend a minimum of two years at Stanford (four years for PhD students) and are expected to undertake significant research projects for either degree. Research opportunities abound, however, and they of course include the several Stanford AIM projects as well as research in psychological and formal statistical approaches to medical decision making, applied instrumentation, large medical databases, and a variety of other applications projects at the medical center and on the main campus. Several students are already contributing in major ways to the AIM projects and core research described elsewhere in this annual report.

We are pleased that the program already has an excellent reputation and is attracting superb candidates for training positions. The program's visibility and reputation is due to a number of factors:

- high quality students, many of whom publish their work in conference proceedings and refereed journals even before receiving their degrees; Stanford MIS students have won first prize in the student paper competition at the Symposium on Computer Applications in Medical Care (SCAMC) in 1985 and 1986, and have also received awards for their work at annual meetings of organizations such as the Society for Medical Decision Making, the American Association for Medical Systems and Informatics (AAMSI), and the American Association for Artificial Intelligence (AAAI);
- a rigorous curriculum that includes newly-developed course offerings that are available to the University's medical students, undergraduates, and computer science students as well as to the program's trainees;
- excellent computing facilities combined with ample and diverse opportunities for medical computer science and medical decision science research;
- the program's great potential for a beneficial impact upon health care delivery in the highly technologic but cost-sensitive era that lies ahead.

The program has been successful in raising financial and equipment support from industry and foundations. It is also recipient of a training grant from the National Library of Medicine. The latter grant was recently renewed for another five years with a study section review that praised both the training and the positive contribution of the SUMEX-AIM environment.

III.A.3.8. Resource Operations and Usage

1 - Operations and Support

The diverse computing environment that SUMEX-AIM provides requires a significant effort at operations and support to keep the resource responsive to community project needs. This includes the planning and management of physical facilities such as machine rooms and communications, system operations routine to backup and retrieve user files in a timely manner, and user support for communications, systems, and software advice. Of course, the move of our groups to new space in the Medical School Office Building has required major planning and care to ensure minimum downtime for our computing environment and much systems and electronics work to outfit the new space.

Our active participation in the planning of the SUMEX/MCS facility in the MSOB resulted in a coordinated environment for twenty-three staff members and thirty-five student workstations, and included 1000 sq. ft. of computer room space and three conference areas. Provisions were made for easily adding more equipment and networking support. The close interaction with the building designers had the additional effect of increasing the designers' interest and knowledge about planning for computer equipment and networking. We have already seen our insight spread to other building projects on campus and the architectural firms will quite likely spread the insight further. Building design appears to be very much an implementation of standards. We have had a part in moving towards the development of more modern standards; certainly here on the campus and perhaps elsewhere.

We use students for much of our operations and related systems programming work. We spend significant time on new product review and evaluation such as Lisp workstations, terminals, communications equipment, network equipment, microprocessor systems, mainframe developments, and peripheral equipment. We also pay close attention to available video production and projection equipment, which has proved so useful in our dissemination efforts involving video tapes of our work.

SUMEX continues to operate with a generally unattended machine room. Our primary operations staff consists of three part-time student workers. This provides a cost-effective approach and gives these undergraduate students an opportunity to participate in the SUMEX project. The major use of this staff is for moving data files to off-line media and to provide data file backup in case of equipment failure. Though we have had nothing that could be classified as a catastrophic failure in the four years of operating our current 2060 equipment, we have had several failures of drives on the SAFE file server. There have been two cases of "soft" failures of disks on the 2060 system. Though these incidents have consumed substantial staff time to deal with, they have not involved significant time loss to the users.

2 - Resource Usage Details

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

1. Overall resource loading data (page 81).
2. Relative system loading by community (page 82).
3. Individual project and community usage (page 85).
4. Network usage data (page 90).

5. System reliability data (page 92).

For the most part, the data used for these plots cover the entire span of the SUMEX-AIM project. This includes data from both the KI-TENEX system and the current DECsystem 2060. At the point where the SUMEX-AIM community switched over to the 2060 (February, 1983), you will notice sharp changes in most of the graphs. This is due to differences in scheduling, accounting, and processor speed calculations between the systems.

2.1 - Overall Resource Loading Data

The following plot displays total CPU time delivered per month. This data includes usage of the KI-TENEX system and the current DECsystem 2060.

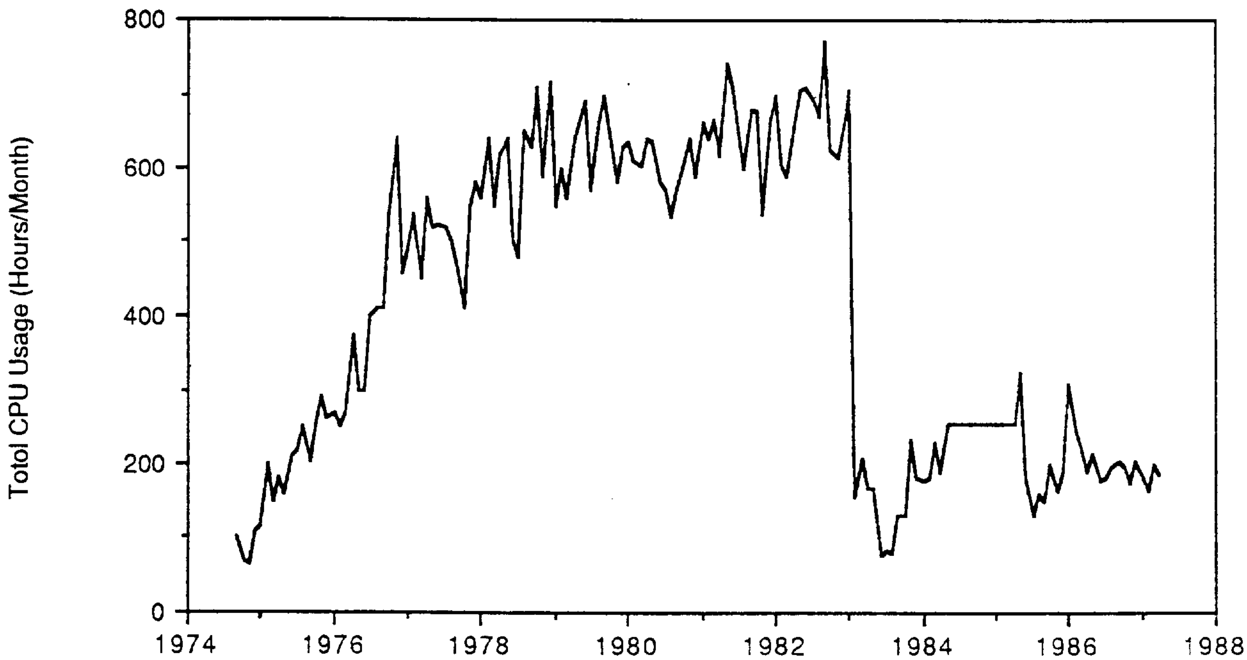


Figure 9: Total CPU Time Consumed by Month

2.2 - 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 nominally divided between these communities as follows:

Stanford	40%
AIM	40%
Staff	20%

The "available" resources to be divided up between these communities 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 80, these plots include both KI-10 and 2060 usage data.

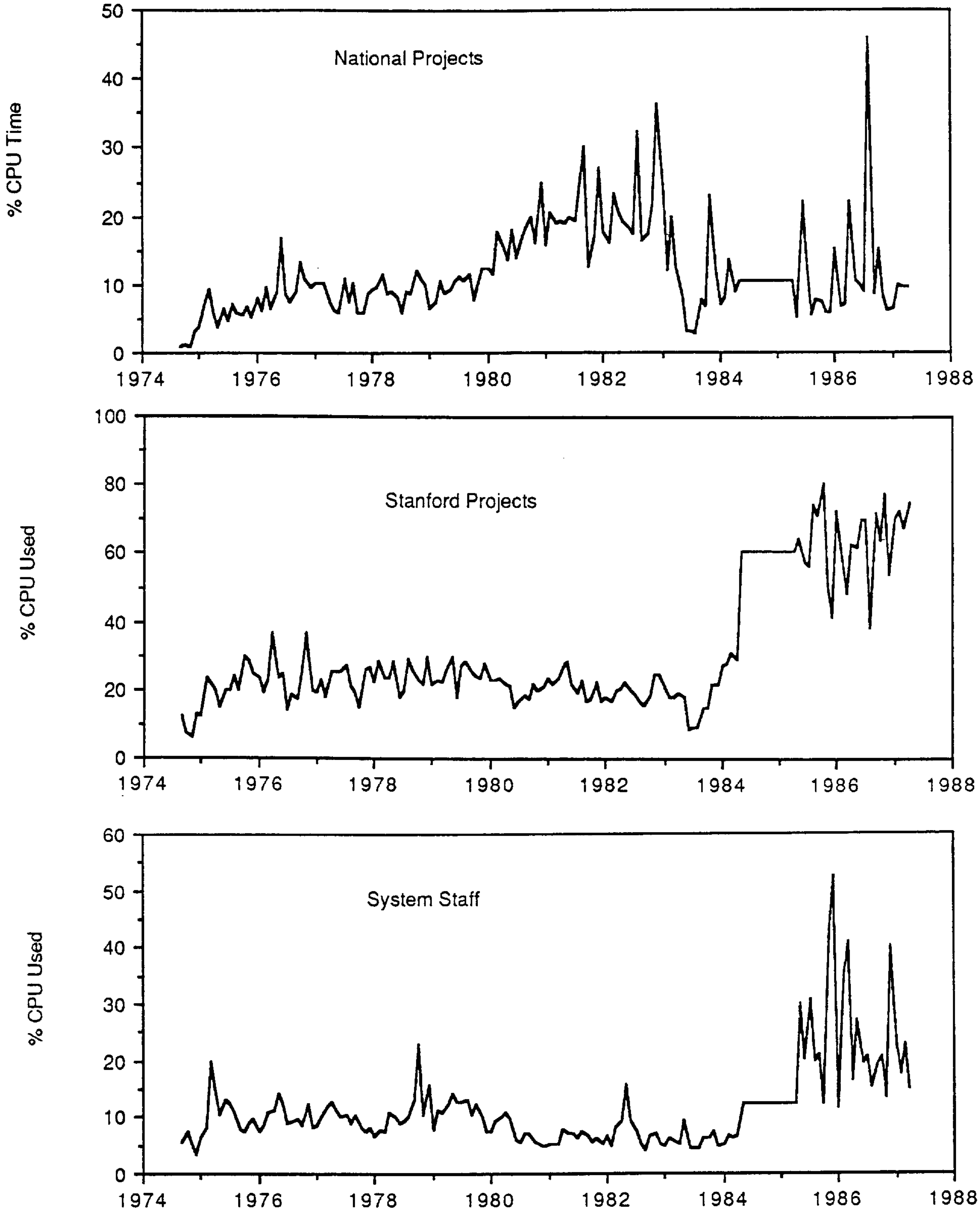


Figure 10: Monthly CPU Usage by Community

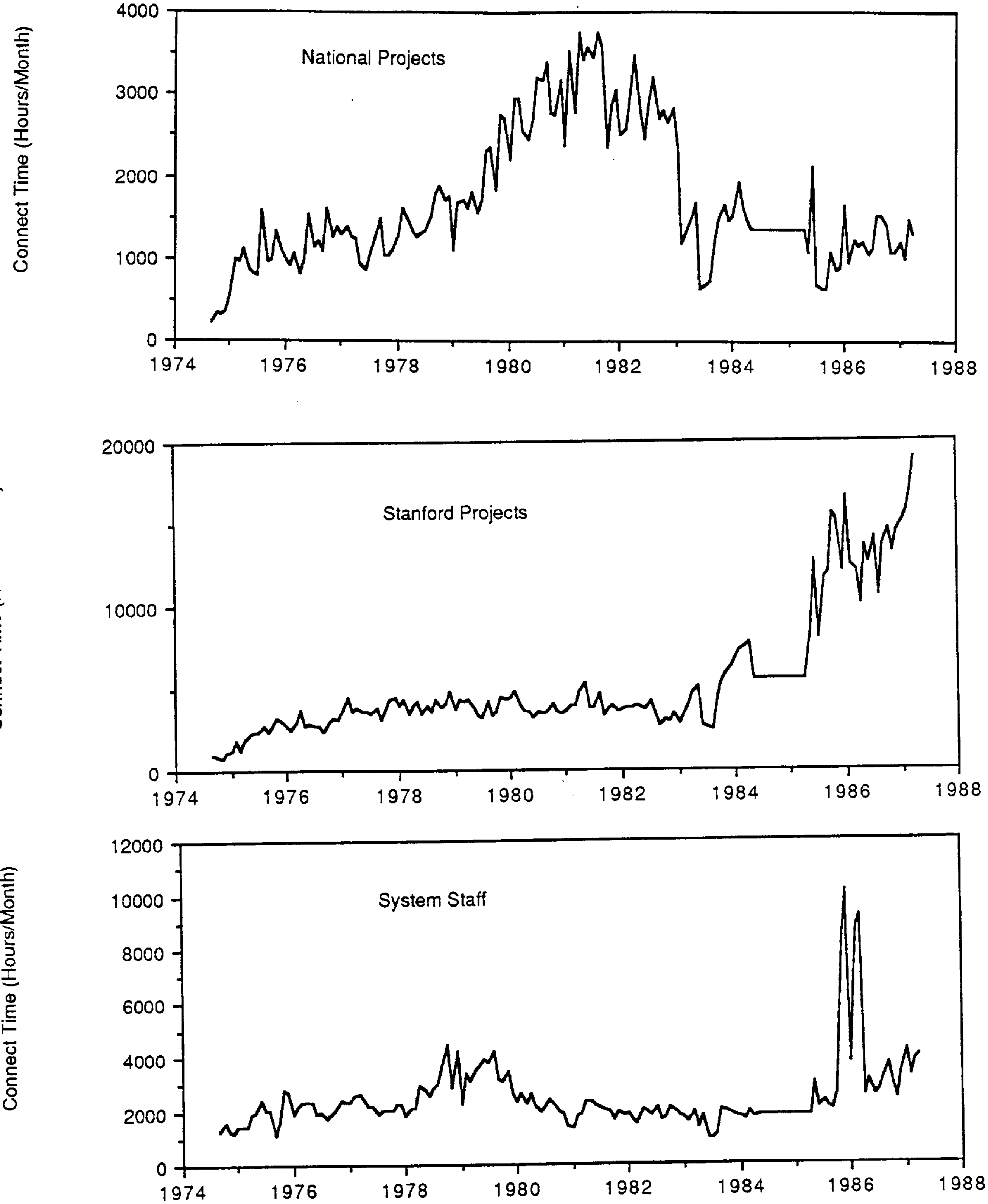


Figure 11: Monthly Terminal Connect Time by Community

2.3 - 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, 1986 and April 30, 1987, on the SUMEX-AIM DECsystem 2060 system.

In the table following, entries include 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 1986 and April 1987.

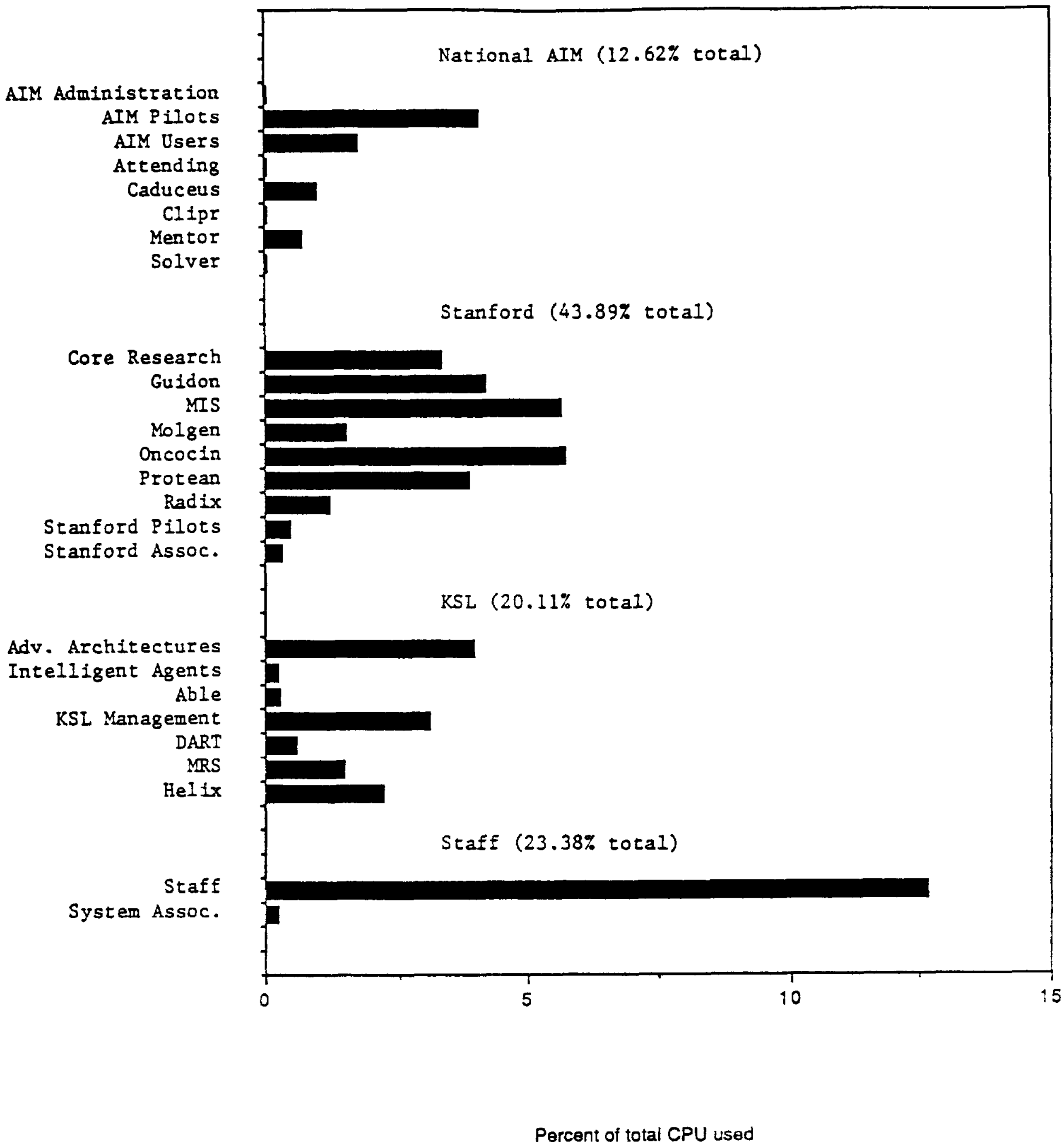


Figure 12: Cumulative CPU Usage Histogram by Project and Community

Resource Use by Individual Project - 5/86 through 4/87

National AIM Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) CADUCEUS "Clinical Decision Systems Research Resource" Jack D. Myers, M.D. Harry E. Pople, Jr., Ph.D. Randolph A. Miller, M.D. University of Pittsburgh	21.63	562.85	1066
2) CLIPR Project "Hierarchical Models of Human Cognition" Walter Kintsch, Ph.D. Peter G. Polson, Ph.D. University of Colorado	0.56	144.73	176
3) SOLVER Project "Problem Solving Expertise" Paul E. Johnson, Ph.D. William B. Thompson, Ph.D. University of Minnesota	0.93	133.87	567
4) MENTOR Project "Medical Evaluation of Therapeutic Orders" Stuart M. Speedie, Ph.D. University of Maryland Terrence F. Blaschke, M.D. Stanford University	16.06	6607.36	1044
5) ATTENDING "A Critiquing Approach to Expert Computer Advice" Perry L. Miller, M.D., Ph.D. Yale University School of Medicine	0.16	24.91	3
6) AIM Pilot Projects	92.49	2813.31	836
7) AIM Administration	0.15	23.82	172
8) AIM Users	40.22	4640.05	2308
	---------	----------	-------
Community Totals	172.27	14967.35	6073

Stanford Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) GUIDON-NEOMYCIN Project Bruce G. Buchanan, Ph.D. William J. Clancey, Ph.D. Dept. Computer Science	95.67	10725.08	1980
2) MOLGEN Project "Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing and Modification" Edward A. Feigenbaum, Ph.D. Peter Friedland, Ph.D. Charles Yanofsky, Ph.D. Depts. Computer Science/ Biology	34.46	7540.02	3109
3) ONCOCIN Project "Knowledge Engineering for Med. Consultation" Edward H. Shortliffe, M.D., Ph.D. Dept. Medicine	131.04	24884.82	2871
4) PROTEAN PROJECT Oleg Jardetzky School of Medicine Bruce Buchanan Computer Science Department	88.64	12876.18	3050
5) RADIX Project Robert L. Blum, M.D. Gio C.M. Wiederhold, Ph.D. Depts. Computer Science/ Medicine	27.01	4356.19	828
6) Stanford Pilot Projects	10.48	1368.01	1690
7) Core AI Research	76.10	16019.67	2732
8) Stanford Associates	6.74	2421.81	179
9) Medical Information Sciences	129.09	16857.16	2060
	---------	----------	-------
Community Totals	599.24	97048.94	18499

KSL-AI Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Advanced Architectures	90.41	3643.33	2999
2) FOL	21.80	2275.61	0
3) Intelligent Agent	5.01	824.73	720
4) KSL Administration	70.26	13298.97	2897
5) DART	12.99	3227.10	1577
6) MRS	33.54	9642.75	2205
7) Helix	50.72	13664.49	802
8) ABLE	6.12	3643.33	233
	---------	----------	-------
Community totals	274.49	74378.94	12512

SUMEX Staff	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Staff	288.63	38743.11	12992
2) System Associates	5.41	681.22	471
	---------	----------	-------
Community Totals	319.20	42391.73	13471

System Operations	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Operations	918.73	95782.32	24016
	=========	==========	=======
Resource Totals	2283.94	324569.30	74571

2.4 - Network Usage Statistics

The plots in Figure 13 and Figure 14 show the monthly network terminal connect time for the public data networks 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.

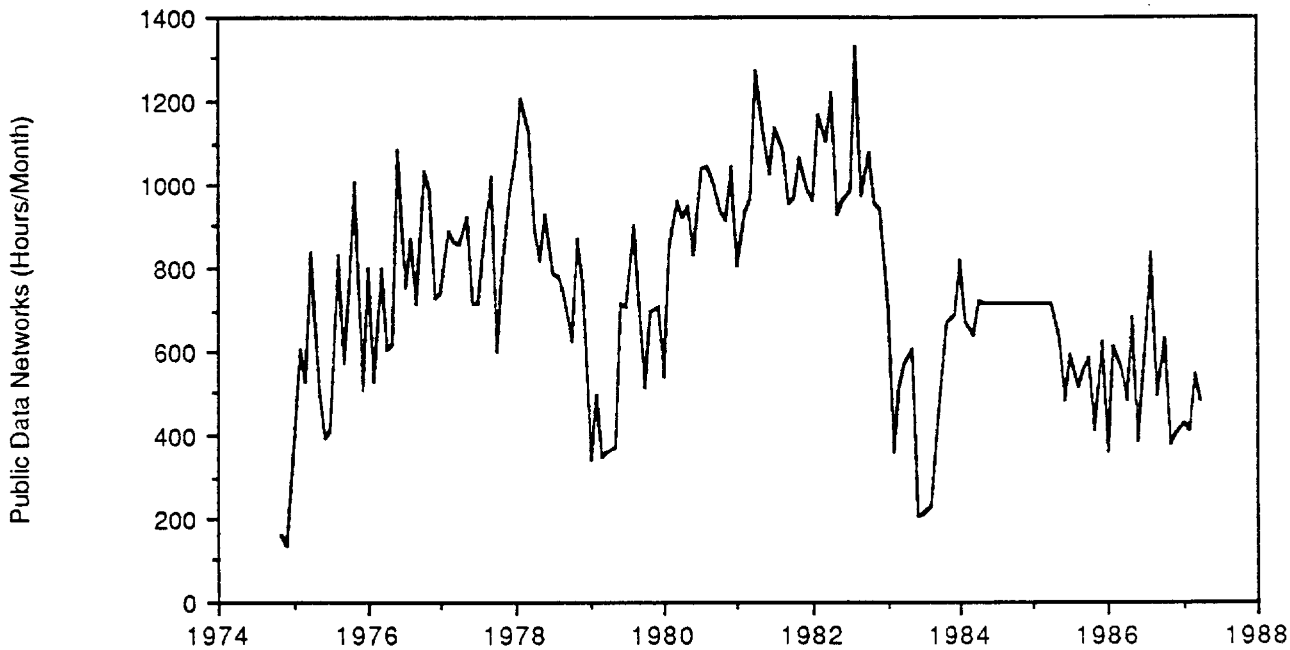


Figure 13: Public Data Network Terminal Connect Time

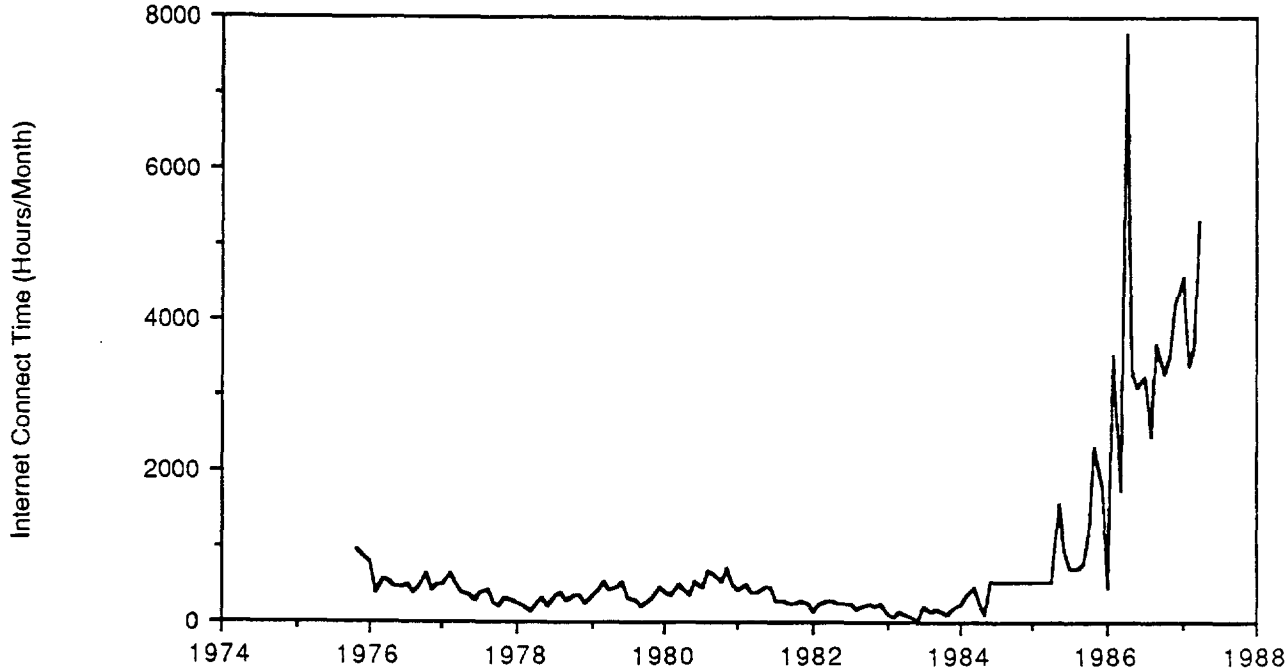


Figure 14: INTERNET Terminal Connect Time

2.5 - System Reliability

System reliability for the DECsystem 2060 has remained quite high in this past period. We have had very few periods of particular hardware or software problems other than while tracking down the internet free space software bug. The data below covers the period of May 1, 1986 to April 30, 1987. The actual downtime was rounded to the nearest hour.

May 1986 - April 1987:

May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
10	28	13	3	2	2	20	2	2	1	9	11

Figure 15: System Downtime -- Hours per Month

May 1986 - April 1987:

Reporting period	:	365 days, 0 hours, 12 minutes, and 49 seconds
Total Up Time	:	359 days, 23 hours, 8 minutes, and 11 seconds
PM Downtime	:	0 days, 18 hours, 35 minutes, and 5 seconds
Actual Downtime	:	4 days, 6 hours, 29 minutes, and 33 seconds
Total Downtime	:	5 days, 1 hour, 4 minutes, and 38 seconds
Mtbf	:	2 days, 13 hours, 42 minutes, and 29 seconds
Uptime Percentage	:	98.83

Figure 16: Overall System Reliability Summary

III.B. Highlights

In this section we describe several research highlights from the past year's activities. These include notes on existing projects that have passed important milestones, new pilot projects that have shown progress in their initial stages, and other core research and special activities that reflect the progress, impact, and influence the SUMEX-AIM resource has had in the scientific and educational communities.

III.B.1. The MENTOR Project

The MENTOR (Medical EvaluatioN of Therapeutic ORders) project, under Drs. Terrence Blaschke at Stanford University, Stuart Speedie at the University of Maryland, and Charles Friedman at the University of North Carolina, seeks to design and develop an expert system for monitoring drug therapy for hospitalized patients. The purpose of the system is to provide appropriate advice to physicians concerning the existence and management of adverse drug reactions.

The computer as a record-keeping device is becoming increasingly common in hospital health care, but much of its potential remains unrealized. Often, information is provided to the physician in the form of raw data. The wealth of such data may effectively hide important information about the patient. 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 using sophisticated medical knowledge and problem solving. Expert systems offer the possibility of embedding this expertise in a computer system which would automatically gather the appropriate information and monitor for the prospect or actual occurrence of adverse drug reactions.

The MENTOR project was initiated in December 1983. The project has been funded by the National Center for Health Services Research since January 1, 1985. As of June 1, 1987, a working prototype system has been developed and is undergoing evaluation. The prototype consists of a Patient Data Base, an Inference Engine, an Advisory Module, and a Medical Knowledge Base. The Medical Knowledge Base currently contains information related to Aminoglycoside Therapy, Digoxin therapy, Surgical Prophylaxis, and Microbiology Lab reports. The system is currently implemented on a Xerox workstation. Another version of the Patient Data Base has been developed for a mainframe and is currently being tested. Plans call for the interconnection of the mainframe and the workstation running the inference engine. The mainframe will then be connected to a Hospital Information System for data acquisition.

III.B.2. The GUIDON Project

The GUIDON/NEOMYCIN Project, under Drs. William J. Clancey and. Bruce G. Buchanan of Stanford University, is a research program to develop a knowledge-based tutoring system for application to medicine. The primary goal 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. The problem-solving procedures are developed by running test cases through NEOMYCIN and comparing them to expert behavior. Also, we use NEOMYCIN as a test bed for the explanation capabilities incorporated in our instructional programs.

Our current emphasis is to construct a knowledge-based 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 has 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 that they can be used with many medical (and non-medical) domains.

It has long been felt that medical knowledge, initially codified for the purpose of computer-assisted consultations, may also be used to teach medical students. The technical basis of the system has matured enough that we are now collaborating closely with medical students and physicians to design a useful tutoring program. The system implements a three-step tutorial process in which the student will solve a problem, watch the system solve it, and then explain his solution and seek explanations about the system's solution. In this way, the program will serve as a *model* that the student can study and compare to his own reasoning.

Another tutorial project involves development of a modeling program (ODYSSEUS) aimed at discovering discrepancies between an expert system knowledge base and that of a student or expert problem solver. When ODYSSEUS watches a student, it functions as a student modeling program and when it watches an expert, it functions as a knowledge acquisition program.

The final major effort involves generalizing our expert system tool, HERACLES, so that it can be made available to other research groups wishing to develop knowledge bases that can be used for tutoring.

In our current work, we are focusing on the modeling, explanation, and knowledge acquisition capabilities that will allow the tutor to articulate how a diagnostic solution is flawed and how it can be improved using specific domain knowledge. Thus, we are teaching the students what constraints a good solution must respect and giving them a language for articulating which medical facts are relevant to the case at hand.

Physicians have generally been enthusiastic about the potential of these programs and what they reveal about current approaches to computer-based medical decision making.

III.B.3. The PROTEAN Project

The PROTEAN project, under Professors Oleg Jardetzky and Bruce Buchanan at Stanford University, is concerned with using artificial intelligence methods to aid in the determination of the 3-dimensional structure of proteins in solution (as opposed to crystallized proteins). 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 allow the study of proteins whose structure cannot be determined with other techniques, and will decrease the time needed for the determination. It is hoped that empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure.

During the past year, we have extended our initial prototype program, called PROTEAN, designed using a *blackboard model*. It is implemented in BB1, a framework system for building blackboard systems that control their own problem-solving behavior. The reasoning component of PROTEAN directs the actions of the Geometry System (GS), a set of programs that performs the computationally intensive task of positioning portions of a molecule with respect to each other in three dimensions. The GS runs in the UNIX environment on a Silicon Graphics IRIS 3020 graphics workstation. The reasoning program (in LISP in BB1) is coupled to the GS by a local area computer network developed by SUMEX.

Pictures of the results of GS computations are displayed on the graphics screen of the IRIS workstation, using a locally developed program called DISPLAY to draw the evolving protein structures at several levels of detail. The DISPLAY program can be used to view structures generated by the GS either under the direct control of the user or as directed by the reasoning system running in BB1. MIDAS and MMS are two other molecular modeling and display systems to manipulate protein structures, particularly those obtained from crystallographic techniques as found in the Protein Data Bank. The ability to observe structures in three dimensions is essential to understanding the behavior of the PROTEAN's reasoning and geometry systems and provides essential insights on the problem solving process.

In addition to the Lac-repressor headpiece protein, we have applied PROTEAN to sperm whale myoglobin, T4 lysozyme, and cytochrome B. Each of these latter proteins has a known crystal structure. In each case, we extracted features of the protein and distance constraints to build data sets for PROTEAN. We then applied the PROTEAN system to the resulting data sets to determine the behavior of the system with different kinds of input.

To determine the correctness and capabilities of the PROTEAN method, we applied PROTEAN to sperm whale myoglobin, a molecule whose crystal structure is known. We systematically explored the dependence of the precision and accuracy of the solutions on the quality of the input data available. In all cases, the sets solutions obtained from PROTEAN include the actual structure of the molecule, with the best results coming from data representing many short range constraints.

Work is proceeding on several aspects of the protein structure problem, including assembly of several partial arrangements and integration of these pieces of solution into larger structures, using atomic level volume exclusion of atoms and information on sidechain packing to produce more precise atomic level solutions, and developing more appropriate representations for unstructured coil sections of proteins.

III.B.4. The Medical Information Science Program

The *Medical Information Sciences (MIS)* program continues to be one of the most obvious signs of the academic impact of the SUMEX-AIM resource on Stanford University. The MIS program received recent University approval (in October 1982) as an innovative training program that offers MS and PhD degrees to individuals with a career commitment to applying computers and decision sciences in the field of medicine. In Spring 1987, a University-appointed review group unanimously recommended that the degree program be continued for another five years. The MIS training program is based in the School of Medicine, directed by Dr. Shortliffe, co-directed by Dr. Fagan, and overseen by a group of six University faculty that includes two faculty from the Knowledge Systems Laboratory (Profs. Shortliffe and Buchanan). It was Stanford's active on-going research in medical computer science, plus a world-wide reputation for the excellence and rigor of those research efforts, that persuaded the University that the field warranted a new academic degree program in the area. A group of faculty from the medical school and the computer science department argued that research in medical computing has historically been constrained by a lack of talented individuals who have a solid footing in both the medical and computer science fields. The specialized curriculum offered by the new program is intended to overcome the limitations of previous training options. It focuses on the development of a new generation of researchers with a commitment to developing new knowledge about optimal methods for developing practical computer-based solutions to biomedical needs.

The program accepted its first class of four trainees in the summer of 1983 and has now reached its steady-state size of approximately 20 graduate students. We do not wish to provide too narrow a definition of what kinds of prior training are pertinent because of the interdisciplinary nature of the field. The program has accordingly encouraged applications from any of the following:

- medical students who wish to combine MD training with formal degree work and research experience in MIS;
- physicians who wish to obtain formal MIS training after their MD or their residency, perhaps in conjunction with a clinical fellowship at Stanford Medical Center;
- recent BA or BS graduates who have decided on a career applying computer science in the medical world;
- current Stanford undergraduates who wish to extend their Stanford training an extra year in order to obtain a "co-terminus" MS in the MIS program;
- recent PhD graduates who wish post-doctoral training, perhaps with the formal MS credential, to complement their primary field of training.

In addition, a special one-year MS program is available for established academic medical researchers who may wish to augment their computing and statistical skills during a sabbatical break. As of Spring 1987, half our trainees have previously received MD degrees and another quarter are medical students enrolled in joint degree programs. One-third are candidates for the MS degree, while the rest are doctoral students. The program has three graduates to date, with several more expecting to complete degrees before the end of 1987. Research opportunities for students include the several Stanford AIM projects as well as research in psychological and formal statistical approaches to medical decision making, applied instrumentation, large medical databases, and a variety of other applications projects at the medical center and on the main campus. Several students are already contributing in major ways to the AIM projects and core research described elsewhere in this annual report.

III.B.5. Remote Virtual Graphics

Lisp workstations of various types have proven extremely powerful, both as development environments for artificial intelligence research and as vehicles for disseminating AI systems into user communities. In addition to the compact, inexpensive computing resources workstations provide, high-quality graphics play a key role in their power. Such graphics systems have become indispensable for understanding the complex data structures involved in developing and debugging large AI systems and are important in facilitating user access to working programs (e.g., for ONCOCIN and PROTEAN).

In the past, members of the SUMEX-AIM community have often watched each others' programs work by linking their CRT terminals to the text output of a running program on the SUMEX 2060. With workstations, though, it is much more difficult to connect to a remote machine and be able to view the complex graphics output of a program. One would like to be able to provide the same powerful graphical tools and programming environment that are available to a user sitting in front of the workstation to the remote user if that user has a low-cost bit-mapped display and mouse.

During this past year, we developed a program called TALK to facilitate interactive, electronic communication between users on independent workstations. Layered on the workstation's native editor, the program allows the full use of all editing capabilities in the process of communication, including deletions, corrections and insertions, font changes, underlining, paragraph formatting, etc. Since the workstation's editor also supports both low- and high-level graphics, the program not only facilitates textual exchanges among users, but also allows the sending of screen images (ONCOCIN flow sheet segments, back traces of program error breaks, code fragments, etc.) as well as structured images (which can be modified on the destination workstation and returned), all interactively.

The TALK program allows the use of different user interfaces, the workstation's document editor being just one possibility. We implemented a simpler terminal mode for compatibility with similar programs on other workstations.

The TALK program has been released gradually to increasing numbers of users in order to get feedback and make changes accordingly. The Medical Computer Science group did an extensive test of the system where for a period they used it in place of their normal electronic and non-electronic communication methods whenever possible. This was both a test of the program and an exploration into what people want in the next generation of electronic communication. The TALK program has been released to the Xerox Lisp workstation community as a whole and researchers at Xerox PARC successfully used the program to hold an interactive, graphic, electronic conversation between users at the PARC facility (in California) and Xerox's EuroPARC facility (in England).

III.C. Administrative Changes

There have been few administrative changes within the project this past year. Professor Shortliffe has been on sabbatical at the University of Pennsylvania as projected last year but has stayed in very close contact with SUMEX and the Medical Computer Science group at Stanford through network connections. During this time, Professor Feigenbaum has acted in the formal role of principal investigator. Professor Shortliffe is expected back at Stanford in mid-July.

The move the Medical Computer Science and SUMEX offices into the newly constructed Stanford Medical School Office Building was completed in June 1986. We now occupy approximately 6500 square feet has almost doubled the space available to us. The design of this space has worked out exceedingly well to improve the interactions within our groups.

We have also designed and implemented a cost recovery system as part of phasing out BRTP subsidy of the DEC 2060 facility. The details of this system are discussed on page 101. In summary, we are successfully recovering the projected 20% of 2060 operations costs this year ($71,376) from Stanford users, with the continuing component of NIH support used to protect national users from fees for service, including communications. This additional burden on Stanford projects was absorbed almost entirely in existing direct cost budgets since no supplements were forthcoming from other funding agencies in the middle of on-going grant and contract awards for new computing costs. This has affected staffing and student support directly in our labor-intensive research efforts. All of our new support applications are being written with requests for funds to cover computing charges.

This next year we will increase the cost recovery goal to 40% of projected 2060 operations costs as scheduled in our grant application of June 1985.

III.D. Resource Management and Allocation

III.D.1. Overall Management Plan

Early in the design of the SUMEX-AIM resource, an effective management plan was worked out with the Biotechnology Resources Program (now Biomedical Research Technology Program) at NIH to assure fair administration of the resource for both Stanford and national users and to provide a framework for recruitment and development of a scientifically meritorious community of application projects. This structure has been described in some detail in earlier reports and is documented in our recent renewal application. It has continued to function effectively as summarized below.

- The AIM Executive Committee meets regularly by teleconference to advise on new project applications, discuss resource management policies, plan workshop activities, and conduct other community business. The Advisory Group meets together at the annual AIM workshop to discuss general resource business and individual members are contacted much more frequently to review project applications. (See Appendix A on page 217 for a current listing of AIM committee membership).

- We have actively recruited new application projects and disseminated information about the resource. The number of formal projects in the SUMEX-AIM community still runs at the capacity of our computing resources. With the development of more decentralized computing resources within the AIM community outside of Stanford (see below), the center of mass of our community has naturally shifted toward the growing number of Stanford applications and core research projects. We still, however, actively support new applications in the national community where these are not able to gain access to suitable computing resources on their own.

- With the advice of the Executive Committee, we have awarded pilot project status to promising new application projects and investigators and where appropriate, offered guidance for the more effective formulation of research plans and for the establishment of research collaborations between biomedical and computer science investigators. This past year we have admitted projects under Professors Perry Miller at Yale University, Larry Widman at the University of Texas, Ira Kalet at the University of Washington, and Robert Beck at Dartmouth University. The latter two sought access primarily for communication with the AIM community as they have research computing resources of their own.

- We have carefully reviewed on-going projects with our management committees to maintain a high scientific quality and relevance to our biomedical AI goals and to maximize the resources available for newly developing applications projects. Several fully authorized and pilot projects have been encouraged to develop their own computing resources separate from SUMEX or have been phased off of SUMEX as a result and more productive collaborative ties established for others.

- We continue to provide active support for the AIM workshops. The next one will be held at the University of Washington in conjunction with the American Association for Artificial Intelligence meeting in July 1987. It is being organized jointly by Drs. Ira Kalet of Washington and Larry Fagan of Stanford.

- We have tailored resource policies to aid users whenever possible within our research mandate and available facilities. Our approach to system scheduling, overload control, file space management, etc. all attempt to give users the greatest latitude possible to pursue their research goals consistent with fairly meeting our responsibilities in administering SUMEX as a national resource.

III.D.2. 2060 Cost Center

General Cost Center Structure

Our renewal proposal for the five-year period 8/1/86-7/31/91, submitted to the Division of Research Resources in June 1985, called for phasing out NIH support for DEC 2060 mainframe operations over the course of the grant period and the establishment of a cost center at Stanford to recover the unsubsidized costs of 2060 operations from the user community. This phasing-out process is taking place linearly over five years, with 20% of the 2060 costs being recovered in renewal year 1 (Grant Year 14), 40% in year 2, 60% in year 3, 80% in year 4, and 100% starting in year 5. In this process, we are attempting to minimize the barriers for national projects by using the continuing partial BRTP subsidy to cover their costs for as long as possible. In this past year, use of the 2060 by members of the national AIM community has been free of charge. Thus, the Stanford user projects are bearing the entire brunt of cost recovery during the first few years. Our plan is conservative, however, in that we are doing this gradually and responsibly so that our users can secure the funding resources and make software changes necessary to allow them to relocate to other facilities or move to workstation environments for their research.

To implement this plan, during the summer of 1986, we requested and received approval from the Government Cost and Rate Studies section of Stanford's Controller's Office to establish a 2060 cost center effective August 1, 1986. We set up the cost center with the simplest possible charge structure in order to minimize the accounting and administrative overhead, establishing a charge rate per CPU hour based on our projections of 2060 operations costs and anticipated billable Stanford project CPU usage. The initial rate was established at $95 per CPU hour.

We closely monitored the cost center expenses and revenues during the year. A mid-year analysis of cost center performance indicated that expenses would be somewhat lower and billable CPU usage somewhat higher than originally projected. To produce a year-end (July 31, 1987) break-even condition for the cost center, we lowered the charge rate as of February 1 to $75 per CPU hour. Figure 17 shows the cumulative user revenues collected by month for the period August 1986 through April 1987 as well as the ideal (linear) cost center recovery line.

The cost center rate for Stanford users is expected to increase substantially at the beginning of each succeeding grant year through renewal year 5, as NIH subsidy of 2060 costs is incrementally withdrawn.

Remote Network Costs

Until this year, the costs associated with networking were supported by NIH through Rutgers University. Beginning this grant year, however, NIH is funding our networking costs directly as part of our 2060 operations budget, and we have entered into a contract with TELENET Communications Corporation for networking services. To underscore our commitment to subsidize the national AIM community's 2060 usage as long as possible, we have been paying for TELENET services directly from the SUMEX grant this year on the assumption that national community members would represent the vast majority of TELENET users. However, all other 2060-related expenses are charged directly to the cost center and then charged out to Stanford users according to their CPU usage and to the SUMEX grant in keeping with its level of subsidy of 2060 operations.

This early practice of paying for TELENET services directly from the grant has complicated our accounting procedures, since networking expenses must ultimately be taken into consideration in allocating total annual 2060 operations costs in correct proportions to the resource budget and to Stanford users. Also, a recent analysis of our networking usage indicated that the use of TELENET by Stanford groups is considerably higher than expected. Therefore, since networking services are not being used exclusively by the national user community as originally believed, we plan to change our procedure and charge TELENET costs directly to the cost center in future years.

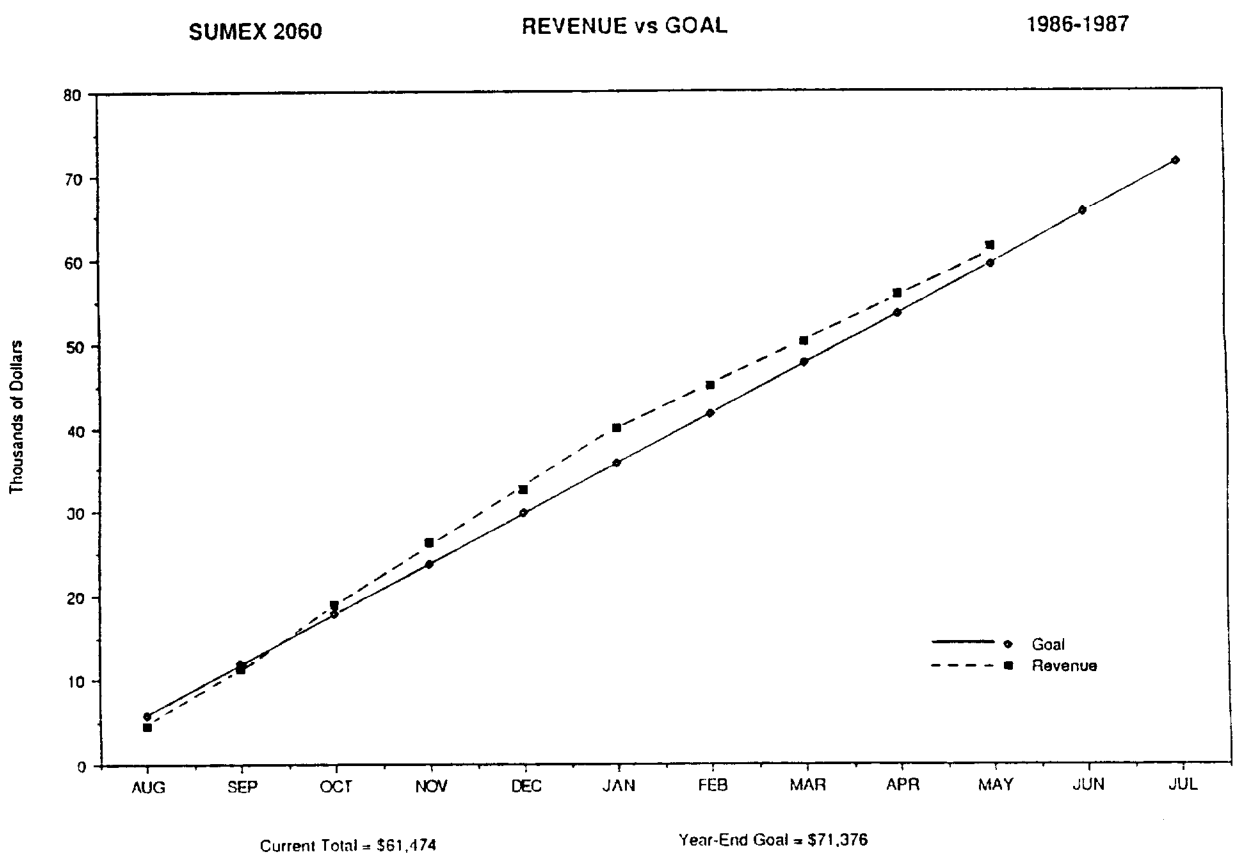


Figure 17: 2060 Cost Center Performance

III.E. Dissemination of Resource Information

We are continuing our past practice of making a substantial effort to disseminate the AI technology developed here. This has taken the form of many publications -- over forty-five combined books and papers are published per year by the KSL; wide distribution of our software including systems software and AI application and tool software, both to other research laboratories and for commercial development; production of films and video tapes depicting aspects of our work; and significant project efforts at studying the dissemination of individual applications systems such as the GENET community (DNA sequence analysis software) and the ONCOCIN resource-related research project (see 123).

Software Distribution

We have widely distributed both our system software and our AI tool software. Since much of our general system-level software is distributed via the ARPANET we do not have complete records of the extent of the distribution. Software such as TOPS-20 monitor enhancements, the Ethernet gateway and TIP programs, the SEAGATE AppleBus to Ethernet gateway, the PUP Leaf server, the SUMACC development system for Macintosh workstations, and our Lisp workstation programs are frequently distributed in this manner to the ARPANET community and beyond. Since our SUMACC software development system for Macintosh workstations is considered to be in the "public domain", we have turned it over to Information Analysis Associates, Mountain View, CA. for distribution (for a minimum charge) to groups not associated with the ARPANET.

Our primary distribution effort is directed towards our AI tool material. In recent years, the volume of inquiries for this type of software and requests for tapes has been a substantial burden on the staff. Records indicate that over the past three years there have been about 1,050 inquiries that have resulted in the distribution of written material about our software systems. It is likely that there have been a similar number of unrecorded or informal interactions on the part of the staff. It was therefore decided to turn over most of this type of software distribution to Stanford's Office of Technology Licensing (OTL).

This organization handles software distribution and technology licensing matters for much of the Stanford community. Since there are several OTL staff members assigned to the distribution of Stanford software, requests for information and tapes are handled quickly and efficiently. Also, OTL's staff has the expertise needed to handle the legal questions that frequently arise in the distribution of software, and an established computerized record-keeping scheme. SUMEX staff continues to be available as needed to assist OTL with special administrative and technical matters.

Unfortunately, start-up delays in the transfer of software distribution to the Office of Technology Licensing and the preparation of new versions of MRS and BB1 have temporarily reduced our distribution volume. During this report period we distributed eleven copies of MRS, eight copies of BB1 and one each of AGE, EMYCIN, GENOA, and CONGEN. During the past year the reconstruction of the distribution packages for the DENDRAL Project (GENOA and CONGEN) has been completed. In December of this year, a five-year exclusive licensing agreement (with Molecular Designs, Ltd.) for the DENDRAL material will expire, and we will therefore have more flexibility in distributing this material.

We continue to make a special effort to assist other members of the SUMEX-AIM community in integrating the technologies needed for biomedical AI research. This is often achieved through direct contact with staff members at these institutions at meetings and workshops or via electronic mailing lists. For example, the Info-1100 mailing list, which is maintained at SUMEX-AIM, has several hundred members (users of Xerox 1100 Series equipment) and is monitored by our staff. This list is used to distribute things like hardware and software bug reports and fixes and system tools and is very valuable to the AIM community Interlisp users.

Video Tapes and Films

The KSL and the ONCOCIN project have prepared several video tapes that provide an overview of the research and research methodologies underlying our work and that demonstrate the capabilities of particular systems. These tapes are available through our groups, the Fleischmann Learning Center at the Stanford Medical Center, and the Stanford Computer Forum, and copies have been mailed to program offices of our various funding sponsors. The three tapes include:

- *Knowledge Engineering in the Heuristic Programming Project* -- This 20-minute film/tape illustrates key ideas in knowledge-based system design and implementation, using examples from ONCOCIN, PROTEAN, and knowledge-based VLSI design systems. It describes the research environment of the KSL and lays out the methodologies of our work and the long-term research goals that guide it.

- *ONCOCIN Overview* -- This is a 30-minute tape providing an overview of the ONCOCIN project. It gives an historical context for the work, discusses the clinical problem and the setting in which the prototype system is being used, and outlines the plans for transferring the system to run on single-user workstations. Brief illustrations of the graphics capabilities of ONCOCIN on a Lisp workstation are also provided.

- *ONCOCIN Demonstration* -- This 1-hour tape provides detailed examples of the key components of the ONCOCIN system. It begins with a demonstration of the prototype system's performance on a time-shared mainframe computer and then shows each of the elements involved in transferring the system to Lisp workstations.

III.F. Suggestions and Comments

Resource Organization

We continue to believe that the Biomedical Research Technology 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 the appropriate technological and medical disciplines.

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 distributed resources of the biomedical community. We would like to see the BRTP take a more active role in promoting these tools within the NIH and its grantee community.

IV. Description of 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 221.

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.

IV.A. 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.

IV.A.1. 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. 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, is 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 use NEOMYCIN as a test bed for the explanation capabilities incorporated in our instructional programs.

The purpose of the current contracts is to construct a knowledge-based 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 has 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 that they 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 used in additional ways that are medically relevant. Using the knowledge to teach medical students is perhaps foremost among these, and GUIDON2 focuses 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 are now collaborating closely with medical students and physicians to design an effective, useful tutoring program. In particular, medical students have served as research assistants, and a current MSAI student is an experienced physician, John Sotos, from Johns Hopkins. The project also collaborates with a community of researchers focusing on medical education, funded by the Josiah Macy, Jr. Foundation.

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

C.1.1 The GUIDON-DEBUG Tutoring Program

We began 1986 with a concerted effort to construct a tutorial program called GUIDON-DEBUG. The idea behind this system is to have a student debug a faulty knowledge base by using graphic explanation and editing tools. A prototype was demonstrated at the annual ONR conference in March. However, after trials with medical students we realized that 1) it was difficult to choose a fault at the right level of difficulty for a student, and 2) the program lacked ability to help the students and evaluate their debugging because it lacked an internal model of how to debug. We concluded that GUIDON-DEBUG development should be deferred until the proposed knowledge acquisition module (see below) is completed.

C.1.2 The GUIDON-MANAGE Tutoring Program

At this point we returned to an alternative conception described in our original proposal, a program called GUIDON-MANAGE. This program teaches a student the language of diagnosis by having him or her enter all requests for patient information as an *abstraction.* Thus, the student issues "strategic commands" such as "test the hypothesis meningitis" or "ask a follow-up question about the headache," and the program (NEOMYCIN) carries out the tactics. By year end, this program was well along, with a complex interpreter for simulating NEOMYCIN to generate help, a feedback window to indicate what NEOMYCIN did when it carried out the commands, and many menus for making input to the program convenient. Research continues to focus on the assistance and feedback components of the program.

GUIDON-MANAGE is now conceived to be the first step in a three-step tutorial program which will include GUIDON-WATCH (which we previously developed) and a yet to be named tutorial module. In these three steps, the student will solve a problem, watch NEOMYCIN solve a problem, and then explain his solution and seek explanations about NEOMYCIN's solution. In this way, we use the program as a *model* that the student can study and compare to his own reasoning.

C.1.3 The GUIDON-MANAGE Tutoring Program

Research in explanation is another major area. This year we completed some difficult programming that allows us to examine a history in detail of everything NEOMYCIN did when solving a problem. With this foundation, we can now go back and summarize lines of reasoning for any point during the previous consultation. In our first program, completed in 1984, we "translated" steps (metarules) using text strings built into the program. Now we seek to generate these strings automatically by having the program read the metarules and select statements to mention. This project makes significant contributions to text generation research, a somewhat ignored area of natural language research.

C.1.4 The ODYSSEUS Modeling Program

Our third tutorial-related project involves continued development of a modeling program, ODYSSEUS. The purpose of ODYSSEUS is to discover domain knowledge discrepancies between an application domain knowledge base (e.g. the Neomycin medical knowledge base) and a student or expert problem solver. IMAGE, an earlier modeling program developed in 1982, did not address this problem. The input to ODYSSEUS is the problem solver's patient data requests. When ODYSSEUS watches a student it functions as a student modeling program for GUIDON2 and when it watches an expert it functions as a knowledge acquisition program for HERACLES.

The approach used by ODYSSEUS to detect domain-level discrepancies may be characterized as failure-driven learning by completing explanations. An explanation failure occurs when ODYSSEUS is unable to create a proof tree consisting of instantiated metarules that links an observable student action to a high-level task goal. In creating these proof trees, a top-down simulation first produces a set of plausible high-level goals and updates problem solving state information; then this information is used by a constrained bottom-up generation from the observable action to these high level goals. A explanation failures occurs when no proof tree can be generated for an action and this suggests a domain level discrepancy.

ODYSSEUS resolves this failures in two steps. First, the constraints on proof tree generation are relaxed; this identifies relations in metarules that might be the source of the discrepancy and produces a set of instantiations for each of these relation that are the candidate domain-level discrepancies. Second, a confirmation theory tests these candidate discrepancies for plausibility.

During the last year, ODYSSEUS has been enhanced to operate directly off an arbitrary set of Heracles control metarules; previously the modeling program incorporated knowledge about the particular metarules that were used in Neomycin. This increases the generality and applicability of the program at the cost of a large increase in the search space. Initial validation tests of Odysseus have been conducted and this has revealed that following the strategic reasoning of human problem solvers is crucially dependent on having a very good domain knowledge base. Besides these tests on human problem solvers, a validation methodology called the synthetic agent method has been designed that allows determination of an upper performance bound. During the next year, ODYSSEUS will be completed, integrated with all parts of Guidon including the explanation and Guidon-Manage program, and validated. A case library for the Neomycin domain will be constructed since this plays a crucial role in validation and assessment of the ODYSSEUS approach.

C.1.5 The HERACLES Expert System Shell

The final major effort involves generalizing our expert system tool, HERACLES, so that it can be made available to other research groups who wish to develop knowledge bases which can be tutored by GUIDON2. This project involves a great deal of basic systems programming, including partitioning of files and regrouping of general and knowledge-base-specific constructs. By year end, we were ready to reconfigure a second program built in HERACLES during 1985, called CASTER, to test out the system-building tools developed to date.

A host of smaller projects included:

- Maintenance of our patient library and records of proper program performance.
- Development of a graphics editor for modifying the knowledge base by "correcting" the program's diagnosis of a particular case.
- Development of menu-based knowledge-base retrieval capability. This program constructs menus that bring together details related to some fact the user has just asked a question about.

- More consistent storage and convenient access to "normal values" for patient tests and findings.

- Development of a package for creating, editing, and replaying "scripts" which take the viewer on a tour of some aspect of NEOMYCIN. Useful for documentation, simple lectures, and automatic demonstrations of the program.

C.1.6 Model of Learning

Finally, in a paper described below, we developed a theory of learning by debugging using knowledge of diagnostic strategy and organization of disease knowledge. This theory now forms the foundation for design of GUIDON2. In our current work, we are focusing on the modeling, explanation, and knowledge acquisition capabilities that will allow the tutor to articulate how a diagnostic solution is flawed and how it can be improved using specific domain knowledge. Thus, we are teaching the constraints a good solution must respect, plus giving the students a language for articulating what medical facts are relevant to the case at hand.

C.1.7 Dissemination of results

There were many conferences relating to our work this year. Most notable were the "Tutoring system workshop" in Windermere, England (travel support from the AAAI) and the "Knowledge acquisition workshop" in Banff, British Columbia. Other useful workshops concerned "Higher-level tools" and "Knowledge compilation." Clancey presented prominent papers at each of these workshops and helped organize the middle two. Clancey also presented Guidon/Neomycin work at additional conferences in Milan, London, New Mexico, Arizona, and Florida.

The Macy Foundation Symposium on Cognitive Science and Medical Education in Montreal, run by John Bruer, was extremely valuable for the grantees. Researchers working on medical instruction included: Feltovich, Evans, Hammond, Elstein, and Patel. Small meetings are unusual in this field (AAAI has more than 5000 attendees); the discussions were detailed and illuminating.

Guidon/Neomycin work will be represented in 1987 at Clancey's tutorial on "Evaluating expert system tools" and his tutorial on tutoring systems at IJCAI in Milan.

C.2 Research in Progress

The following projects are active as of May 1987 (see also near-term plans listed in Section III.A):

1. Developing additional instructional programs based on NEOMYCIN;

2. Studying learning in the setting of debugging a knowledge base;

3. Re-implementing the explanation program to use the logic-encoding of the metarules (stating this program in the same task/metarule language so that it might reason about its own explanations);

4. Developing new graphic methods for making presentations from the knowledge base, including tour-like lectures and "dynamic menus" which bring together items relevant to previous user inquiries;

5. Applying the student modeling program, ODYSSEUS, to knowledge acquisition; and

6. Preparing HERACLES, the generalization of NEOMYCIN, for use by other people.

D. Publications Since January 1986

1. Clancey, W.J., Richer, M., Wilkins, D.C., Barnhouse, S., Kapsner, C., Leserman, D., Macias, J., Merchant, A., and Rodolitz, N.: *Guidon-Debug: The student as knowledge engineer*. KSL Working paper 86-34.

2. Clancey, W.J.: *Qualitative Student Models*. **Annual Review of Computer Science.** Palo Alto: Annual Reviews, Report KSL-86-11, Computer Science Dept., May 1986.

3. Clancey, W.J.: *From GUIDON to NEOMYCIN and HERACLES in twenty short lessons: ONR Final Report, 1979-1985. The AI Magazine*, 7(3):40-60, Conference, 1986.

4. Wilkins, D.C., Clancey, W.J., Buchanan, B.G. An overview of the ODYSSEUS learning apprentice. In **Machine Learning: A Guide to Current** Research, eds. T.M. Mitchell, J.G., Carbonell, and R.S. Michalski. New York, Academic Press, pages 369-373. Also KSL-85-26.

5. Clancey, W.J. Intelligent tutoring systems: A tutorial survey. **International Professorship Series, 1985** Academic Press, Inc., London, in press.

6. Wilkins, D.C., Clancey, W.J., and Buchanan, B.G. On Using and Evaluating Differential Modeling in Intelligent Tutoring and Apprentice Learning Systems. In **Intelligent Tutoring Systems: Lessons Learned**, eds. J. Psotka, D. Massey, and S. Mutter. Lawrence Erlbaum Publishers, in preparation. Also KSL-86-62.

7. Clancey, W.J. The knowledge engineer as student: Metacognitive bases for asking good questions. In **Learning Issues in Intelligent Tutoring** Systems, eds. A. Lesgold and H. Mandl, in preparation. Also KSL 87-12.

8. Clancey, W.J. *Viewing knowledge bases as qualitative models*. KSL Working paper 86-27.

9. Clancey, W.J. *Know-how vs. knowledge representation* (extended abstract). *Proceedings of the Workshop on Knowledge Compilation*, Oregon State Technical report, September 1986, pages 1-2.

10. Wilkins, D.C., Clancey, W.J., and Buchanan, B.G., Knowledge Base Refinement Using Abstract Control Knowledge. January, KSL-87-01.

11. Wilkins, D.C., Buchanan, B.G., and Clancey, W.J., *The Global Credit Assignment Problem and Apprenticeship Learning*. January, KSL-87-04.

12. Clancey, W.J. Review of Winograd and Flores's "Understanding Computers and Cognition": A favorable interpretation. *Artificial Intelligence*, December, 1986.

13. Dietterich, T. G., Flann, N. S. and Wilkins, D. C., Machine Learning at IJCAI-85, in *Machine Learning*, Volume 1, No. 2, 1986, 227-242.

14. Karp, P. D. and Wilkins, D. C., An Analysis of the Deep/Shallow Distinction For Expert Systems, KSL-86-32, April 1986, 18 pp.

15. Wilkins, D. C. and Buchanan, B. G., Debugging Rule Sets When Reasoning Under Uncertainty, in *Proceedings of the Fifth National Conference on Artificial Intelligence*, August 1986, 448-454. Also, extended version, KSL-86-30, 20 pp.

16. Wilkins, D. C.,Knowledge Base Debugging Using Apprenticeship Learning Techniques, in *Proceedings of the Knowledge Acquisition for Knowledge-Based Systems Workshop*, November 1986, 40. 0--40. 14. Also, revised version, KSL-86-63, 20 pp.

17. Wilkins, D. C., Clancey, W. J. and Buchanan, B. J., Knowledge Base Refinement Using Abstract Control Knowledge, to appear in *Knowledge Acquisition for Knowledge Based Systems*, edited by J. Boose and B. Gaines, Academic Press. Also to appear in *International Journal of Man-Machine Studies*. Also KSL-87-01, Dec 1986, 12 pp.

18. Wilkins, D. C., Cognitive Diagnosis of Heuristic Classification Problem Solving, *Third International Conference on Artificial Intelligence and Education*, May 1987, pp 57. Also, KSL-86-71, Dec 1986, 2 pp.

E. Funding Support

Contract Title: "A Family of Intelligent Tutoring Programs for Medical Diagnosis"
Principal Investigator: Bruce G. Buchanan, Prof. Computer Science, Research
Associate Investigator: William J. Clancey, Research Assoc. Computer Science
Agency: Josiah Macy, Jr. Foundation
Term: March 1985 to March 1988
Total award: $503,415 direct costs

Contract Title: "Computer-Based Tutors for Explaining and Managing the Process of Diagnostic Reasoning"
Principal Investigator: Bruce G. Buchanan, Prof. Computer Science, Research
Associate Investigator: William J. Clancey, Research Assoc. Computer Science
Agency: Office of Naval Research
ID number: N00014-85-K-0305
Total award: $510,311 total

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

We are frequently asked to demonstrate GUIDON-MANAGE, GUIDON-WATCH, and NEOMYCIN to Stanford visitors or at meetings in this country or abroad. Physicians have generally been enthusiastic about the potential of these programs and what they reveal about current approaches to computer-based medical decision making. We use network e-mail through SUMEX to communicate with other researchers worldwide.

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 share programming expertise and utility routines. 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

The SUMEX resources group has provided exemplary service. We have no complaints or suggestions whatsoever.

III. RESEARCH PLANS

A. Project Goals and Plans

Research over the next year will continue on several fronts, including one or more prototype instructional programs.

1. Use GUIDON-MANAGE by medical students to empirically develop the interface and teaching scenario.
2. Integrate the new explanation program into the GUIDON-MANAGE program in order to provide explanations of the operations of tasks invoked by the student.
3. Develop the GUIDON-DEBUG knowledge acquisition program and incorporate its perspective on diagnosis (operators for manipulating the patient-specific model) in feedback provided within GUIDON-MANAGE.

B. Long-term plans

Plans beyond 1988 are uncertain at this time. We expect to make HERACLES available for routine use by people outside of Stanford and explore non-medical applications to broaden our understanding of diagnosis and heuristic classification problem solving.

C. Requirements for Continued SUMEX Use

SUMEX remains the central communications facility for our project--for communication by e-mail and for preparing publications. Research is done on SUMEX-supported Lisp workstations.

D. Requirements for Additional Computing Resources

Within eighteen months, we believe that we will need to upgrade existing workstations purchased in the past few years to incorporate new memory sizes and faster processors. Our experience with color monitors on IBM PC's indicates that the research world must convert to color to fully exploit the potential of computer graphics, especially for knowledge base browsing and editing. There is some question whether academic labs will be left behind by industrial efforts in this respect. We also find that the existing printers are unreliable and of uneven quality. These must be replaced in the near future, perhaps at a higher cost for durability.

E. Recommendations for Future Community and Resource Development

With the proliferation of machine types and the availability of stand-alone machines such as the Macintosh, it is important that the machine be linked for convenient communication by e-mail and conventions be established for automatically translating old publication files into new standard formats.

IV.A.2. 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. Our current research concentrates on scientific discovery within the subdomain of regulatory genetics. We desire to 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 two hundred 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 four hundred 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. BIONET is now serving many of the computational needs of over two thousand academic molecular biologists in the United States.

More generally, our work in qualitative simulation as applied to molecular biology is also relevant to building models of many other medical and biological systems. For example, one Artificial Intelligence researcher (Kuipers) has been applying these techniques to the domain of renal physiology. Other researchers within the KSL are considering applying these techniques to building models of cardio-pulmonary physiology.

C. Highlights of Research Progress

C.1 Accomplishments

During the past year we have concentrated on the qualitative modeling and simulation aspects of the research. Our view is that a well-formulated, multi-level model of a scientific theory is a necessary first step to automated discovery. In addition, we have worked on knowledge acquisition and graphical display of process information and on the description and understanding of the results of laboratory experiments. We have also prepared an in-depth conceptual reconstruction of the biological research which led to the current detailed understanding of the mechanism of attenuation. The highlights of this work are summarized in several categories below.

C.1.1 Qualitative Modeling and Simulation

Our work in qualitative simulation has been directed towards building a program which embodies a theory of the tryptophan system. We have built one model of the system and we are designing a second model based on the successes and failures of the first.

The first model is organized around a set of twenty important state variables of the tryptophan system which we have identified. In addition, it contains descriptions of the causal interactions between these state variables. The novel properties of this model results from the novel representations used for the state variables and the interactions between them.

Our approach to the representation of the values of state variables results from two observations. First, the amount of information biologists have about the values of different state variables varies widely. Second, different amounts of information about a given variable may be available, and of interest, for different problems. Thus, our representation is designed to capture a variety of types of statements about the value of a variable. For example, we can record quantitative information about a variable ($x = .05$), inequality information ($x > 10$), or relative information ($x = 2*y$).

Just as there is a range in the degree of precision with which we might know the value of a given variable, there is an analogous range within which we might know the causal relationship between two variables. Consider that there does exist some function which describes the interactions among any set of variables in our system. Biologists may not have been able to determine the exact behavior of this function, and hence cannot describe it exactly. Or, we may know its exact behavior, but it may be so complex that we wish to describe it more simply.

Thus, we require a set of representations which allows us to represent the exact form of a function if we have it, or approximations if we do not have it or it is too complex. Relationships among variables are concepts which are represented with several frames, within which all or only some slots may be filled. Relationships between each pair of interacting variables are represented with frames called *Relations*, which describe a unidirectional causal relationship between two variables. For example, we can record any of:

- the sign of a relationship
- whether it is a monotonic relationship
- what the functional form of the relationship is, e.g., linear, higher polynomial, exponential, or unknown
- the sign of the exponent on the input variable
- one or more quantitative coefficients for the relationship

Using these representations we can thus express precisely that (possibly incomplete) knowledge that biologists have about the trp system. We can then define experimental conditions and ask the simulation system to make predictions as to the degree of expression of the genes in the tryptophan operon. For example, we can ask how much expression occurs when the cell is starved of tryptophan, or when tryptophan is in excess. The simulation system propagates the initial experimental conditions through the model in a cyclic fashion to predict how the expression of the operon varies over time.

C.1.2 Process Description and Graphical Display

A system has been built which generalizes our experience in process description by providing a simplified interface for the domain-independent description and animation of process knowledge. The system allows processes to be broken down into component sub-processes and the causal and time-oriented relationships of the subprocesses to be specified. In addition, objects utilized by the processes can be conveniently described and "drawn" with modes and points of interaction among the objects given by the user. All knowledge about processes and objects is automatically stored in the framework of a KEE knowledge base.

After process and object description, the system automatically animates the process by displaying one of several primitive types of interactions among objects in the proper time order dictated by the process knowledge base. This system has been tested on the tryptophan operon domain and its utility is currently being explored in a medical simulation domain.

C.1.3 A Conceptual Reconstruction of the Discovery of Attenuation

Scientific theory formation is a complicated process. The construction of a computer program to reproduce scientific discoveries is one way to study this process. Another way to study the process is by studying the work of actual scientists.

In the past year we have prepared an in-depth study of the discovery of attenuation by Charles Yanofsky and other researchers. We have studied the biological literature extensively and interviewed many scientists involved in the research in order to reconstruct the different conceptual states of knowledge through which the scientists passed in their understanding of the tryptophan operon. By analyzing these states of knowledge and the transitions between them, we have elucidated a number of the strategies and heuristics which these biologists used to generate and choose between theories of the tryptophan operon. We have related these strategies to both the ideas of different philosophers of science, and to the diagnostic strategies of the Internist medical expert system.

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 Kedes, L.: *Discovering the secrets of DNA.* Communications of the ACM, 28(11):1164-1186, November, 1985, and IEEE/Computer, 18(11):49:69, November, 1985.

12. Friedland, P. and Iwasaki Y.: *The concept and implementation of skeletal plans*. Journal of Automated Reasoning, 1(2): 161-208, 1985.

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

14. 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.

15. 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.

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

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

18. 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.

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

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

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

22. Karp, P., and D. Wilkins: *An Analysis of the Deep/Shallow Distinction for Expert Systems.* Stanford University Knowledge Systems Laboratory Report KSL-86-32, 1986.

23. Karp, P., and P. Friedland: *Coordinating the Use of Qualitative and Quantitative Knowledge in Declarative Device Modeling.* Stanford University Knowledge Systems Laboratory Report KSL-87-09, 1987.

24. Round, A.: *QSOPS: A Workbench Environment for the Qualitative Simulation of Physical Processes.* Stanford University Knowledge Systems Laboratory Report KSL-87-37, 1987.

25. Karp, P., and P. Friedland, *A Conceptual Reconstruction of the Discovery of Attenuation.* In Preparation.

E. Funding Support

The MOLGEN grant, which has supported the bulk of this research, is titled: MOLGEN: Applications of Artificial Intelligence to Molecular Biology: Research in Theory Formation, Testing, and Modification. This NSF Grant number MCS-8310236, expired on 10/31/86. The Principal Investigators were Edward A. Feigenbaum, Professor of Computer Science and Charles Yanofsky, Professor of Biology. Additional support for this research has been provided by the Defense Advanced Research Projects Agency, under contract N00039-86C-0033.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

SUMEX-AIM continues to serve as the nucleus 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 EMACS, MM, Scribe and BULLETIN BOARD have made possible the project's 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 into which SUMEX has evolved. 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 currently have a satisfactory mode of operation where essentially all 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 to colleagues 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. We will continue our work in qualitative simulation, modeling, and process description. We will continue testing the existing state-variable-based model of the tryptophan operon. In addition, we will construct a new and more general model of the operon. This model will be centered around the objects within this domain (e.g., enzymes, DNA, repressor proteins) and the interactions between them. The current state-variable model makes assumptions about the presence of different objects and the functions of these objects (e.g., that they contain no mutations) which the new model will make both explicit and allow us to change. Essentially, the new model will allow us to dynamically construct new state-variable models based on the presence of different objects and different interactions between these objects. Changing these assumptions is crucial to the discovery process, which involves the postulation of new classes of objects and new classes of interactions between objects.

2. Build a mechanism for postulating extensions or corrections to the current theory: a constrained theory generator. Our conceptual reconstruction of the discovery of attenuation should be of critical help in both this phase and the phases which follow.

3. Build a mechanism for evaluating alternative theories. This would include rating the theories based on plausibility, selectability, completeness, significance, and so on. We hope the evaluation process produces information useful in discriminating among the possible theories.

4. Test the entire structure on the evolving trp operon regulatory system. Experiment with different initial knowledge bases to see how the discovery process is altered by the availability of new techniques, analogous systems, and so forth.

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.

IV.A.3. 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 ensure 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

The ONCOCIN Project is a large interdisciplinary effort that has involved over 35 individuals since the project's inception in July 1979. The work is currently in its eighth year; we summarize here the milestones that have occurred in the research to date:

- *Year 1:* The project began with two programmers (Carli Scott and Miriam Bischoff), a Clinical Specialist (Dr. Bruce Campbell) and students under the direction of Dr. Shortliffe and Dr. Charlotte Jacobs from the Division of Oncology. 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.

- *Year 2:* The following year (1980-1981) 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 of the documents 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.

- *Year 3:* 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.

- *Year 4:* 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 became an expert on the Xerox workstations that we are using. 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.

- *Year 5:* The project's fifth year (1983-1984) was characterized by growth in the size of our staff (three new full-time staff members and a new oncologist joined the group). The increased size resulted from a DRR grant that permitted us to begin a major effort to rewrite ONCOCIN to run on professional workstations. Dr. Robert Carlson, who had been our Clinical Specialist for the previous two years, was replaced by Dr. Joel Bernstein, while Dr. Carlson assumed a position with the nearby Northern California Oncology Group; this appointment permitted him to continue his affiliation both with Stanford and with our research group. In August of 1983, Larry Fagan joined the project to take over the duties of the ONCOCIN Project Director while also becoming the Co-Director of the newly formed Medical Information Sciences Program. Dr. Fagan continues to be in charge of the day-to-day efforts of our research. An additional programmer, Jay Ferguson, joined the group in the fall to assist with the effort required to transfer ONCOCIN from SUMEX to the 1108 workstation. A fourth programmer, Joan Differding, joined the staff to work on our protocol acquisition effort (OPAL).

- *Year 6:* During our sixth year (1984-1985) we further increased the size of our programming staff to help in the major workstation conversion effort. The ONCOCIN and OPAL efforts were greatly facilitated by a successful application for an equipment grant from Xerox Corporation. With a total of 15 Xerox LISP machines now available for our group's research, all full-time programmers have dedicated machines, as do several of the senior graduate students working on the project. Christopher Lane took on full-time responsibility for the integration and maintenance of the group's equipment and associated software. Two of our programming staff moved on to jobs in industry (Bischoff and Ferguson) and three new programmers (David Combs, Cliff Wulfman, and Samson Tu) were hired to fill the void created by their departure and by the reassignment of Christopher Lane.

  In addition to funding from DRR for the workstation conversion effort, we have support from the National Library of Medicine which supports our more basic research activities regarding biomedical knowledge representation, knowledge acquisition, therapy planning, and explanation as it relates to the ONCOCIN task domain. We have continued to study the therapy planning process under support from the NLM. This research is led by Dr. Fagan and has concentrated on how to represent the therapy-planning strategies used to decide treatment for patients who run into serious problems while on protocol-described treatment. The physicians who treat these patients often seek out a consultation with the protocol study chairman. Dr. Branimir Sikic, a faculty member from the Stanford University Department of Medicine, and the Study Chairman for the oat cell protocol, collaborated on this project. Janice Rohn joined the ONCOCIN project as data manager and to assist in the knowledge entry process.

- *Year 7:* The seventh year (1985-86) marked several milestones in our research on workstation-based programming. The OPAL knowledge acquisition system became operational, and several new oncology protocols were entered using this system. David Combs was primarily responsible for creating the operational version of OPAL (based on the initial prototype by Joan Differding Walton). As anticipated, we increased the speed and ease with which protocols can be added to the ONCOCIN knowledge base.

  Based on the protocols entered through OPAL, we began experimental testing of the workstation version of ONCOCIN in the Stanford oncology clinic. Clifford Wulfman developed the user interface (based on an initial prototype designed by Christopher Lane). Samson Tu developed the reasoning component (designed originally by Jay Ferguson). Much of their work is built upon an object-oriented system developed for our group by Christopher Lane. We connected the various parts of the system, and demonstrated that we have the capability to run ONCOCIN with the reasoning program and interface program on different machines in the communication network. The current version of the program is currently run on a single workstation, but future versions may take advantage of the multiple machine option. To increase the speed at which we are able to test protocols entered into ONCOCIN, we developed additional programs to test real and synthetic cases without user interaction; these are then reviewed by our collaborating clinicians.

  We also developed a workstation-based program, OPUS, to help clinicians determine which protocols are appropriate for specific patients. OPUS was designed and implemented by Janice Rohn with the assistance of Christopher Lane. We have been using it in the clinic setting since the end of 1985.

Thus, in addition to providing an information resource about protocols, the use of a graphically-oriented program provided a way to learn about the software style and hardware used in the workstation version of ONCOCIN.

We discontinued the mainframe version of ONCOCIN, and began using the workstation version exclusively. The performance of the mainframe version of ONCOCIN was documented in two evaluation papers that appeared in clinical journals (see Hickam and Kent's papers).

We continued our basic research in the design of advanced therapy-planning programs: the ONYX project. We developed a model for planning which includes techniques from the fields of artificial intelligence, simulation, and decision analysis. Artificial intelligence techniques are used to create a small number of possible plans given the ideal therapy and the patient's past treatment history. Simulation techniques and decision analysis are used to examine and order the most promising plans. Our goal is to allow ONCOCIN to give advice in a wider range of situations; in particular, the system should be able to recommend plans for patients who have an unusual response to chemotherapy.

During this year, Stephen Rappaport, M.D. joined us as a programmer on the therapy planning research. Clinical expertise for ONCOCIN was provided by Richard Lenon, M.D. and Robert Carlson, M.D.

- *Year 8:* This year (1986-87) concentrated on two diverse tasks: 1) scaling up the use of the workstation version of ONCOCIN in the clinic, and 2) generalization of each of the components. The latter task is described in the core research sections of this report(see page 19).

  In 1986, we placed the workstation version of ONCOCIN into the Oncology Day Care clinic. This version is a completely different program from the version of ONCOCIN that ran on the DECsystem 20--using protocols entered through the OPAL program, with a new graphical data entry interface, and a revised knowledge representation and reasoning component. One of the Oncology Clinical Fellows (Andy Zelenetz) became responsible for verifying how well our design goals for ONCOCIN had been accomplished. His suggestions have included the addition of key protocols and the ability to have the program used as a data management tool if the complete treatment protocol had not yet been entered into the system. Both of these suggestions were carried out during this year, and the program has achieved wider use in the clinic setting. In addition, laser-printed flowsheets and progress notes have been added to the clinic system.

  The process of entering a large number of treatment protocols in a short period of time led to other research topics including: design of an automated system for producing meaningful test cases for each knowledge base, modification of the design and access methods for the time-oriented database, and the development of methods for graphically viewing multiple protocols that are combined into one large knowledge base. These research efforts will continue into the next year. In addition, some of the treatment regimens developed for the original mainframe version are still in use and can be transferred to the new version of ONCOCIN. The process of converting this knowledge will also be undertaken in the next year. As the knowledge base grows, additional mechanisms will be needed for the incremental update and retraction of protocols. Additional changes in the reasoning and interface components of the system are described below.

A new research project related to ONCOCIN was started this last year. We are exploring the use of continuous speech recognition as an alternate entry method for communicating with ONCOCIN. This project requires the connection of speech recognition equipment produced by Speech Systems, Inc. of Tarzana to the ONCOCIN interface module. Christopher Lane has already developed a prototype network connection and command interpreter between the speech module (running on a Sun with special hardware added) and the Xerox 1186 computer that runs ONCOCIN. Clifford Wulfman has designed a series of modifications to the ONCOCIN user interface to allow for verbal commands. Graduate student Danielle Fafchamps has helped to design experiments to elicit how clinicians would like to phrase their requests to ONCOCIN.

Janice Rohn is creating a new version of the Librarian program which facilitates the physician's initial communication with the ONCOCIN system (based on the original version by Cliff Wulfman). We continue to collaborate with Andy Zelenetz, Richard Lenon, Robert Carlson, and Charlotte Jacobs on the design and implementation of ONCOCIN in the clinic. Stephen Rappaport has started a residency program to continue his medical education.

C.2 Research in Progress

Our research in the ONCOCIN project over the last year comprised three major categories: (1) conversion of ONCOCIN to the workstation version, (2) development of a knowledge acquisition interface (OPAL) for entering new protocols, and (3) modeling of the strategic therapy selection process (ONYX). We are now able to explore ways to test the system beyond the Stanford environment.

A summary of our current research endeavors follows.

C.2.1 Transfer of the ONCOCIN system from the DEC-20 to the Xerox 1100 Series machines

During the process of converting to the workstation version of ONCOCIN, we redesigned segments of the program. We have completed the major portion of that work, and our experience with the new version has suggested additional areas for improving the reasoning techniques and knowledge representation of ONCOCIN.

- *Redesign of the reasoning component.* A major impetus for the redesign of the system was to develop more efficient methods to search the knowledge base during the running of a case. We have implemented a reasoning program that uses a discrimination network to process the cancer protocols. This network provides for a compact representation of information which is common to many protocols but does not require the program to consider and then disregard information related to protocols that are irrelevant to a particular patient. We continue to improve portions of the reasoning component that are associated with reasoning over time; e.g., modeling the appropriate timing for ordering tests and identifying the information which needs to be gathered before the next clinic visit. In general, we are concentrating on improving the representation of the knowledge regarding sequences of therapy actions specified by the protocol.

  Our experience with adding a large number of protocols has led to the evaluation of the design of the internal structure of the knowledge base (e.g., the way we describe the relationships between chemotherapies, drugs, and treatment visits). We will continue to improve the method for traversing the plan structure in the knowledge base, and consider alternative arrangements for representing the structure of chemotherapy plans. Currently, the knowledge base of treatment guidelines and the patient database are separated. We propose to tie these two structures closer together. Additional work is anticipated on turning ONCOCIN into a critiquing system, where the physician enters their therapy and ONCOCIN provides suggestions about possible alternatives to the entered therapy. Although we have concentrated our review of the ONCOCIN design primarily on the data provided by additional protocols, we know that non-cancer therapy problems may also raise similar issues. The E-ONCOCIN effort is designed to produce a domain-independent therapy planning system that includes the lessons learned from our oncology research. Samson Tu is primarily responsible for continued improvement of the reasoning component of ONCOCIN.

- *Development of a temporal network.* The ability to represent temporal information is a key element of programs that must reason about treatment protocols. The earlier version of the ONCOCIN system did not have an explicit structure for reasoning about time-oriented events. We are experimenting with different configurations of the temporal network, and with the syntax for querying the network. We are also adapting this network so that it can interface with the ONYX therapy-planning systems. This research on temporal reasoning is part of Michael Kahn's Ph.D. thesis. Michael is a student in the Medical Information Sciences Program at University of California at San Francisco.

- *Extensions to the user interface.* We continue to experiment with various configurations of the user interface. Many of the changes have been in response to requests for a more flexible data management environment. We are occasionally faced with data that becomes available corresponding to a time before the current visit. This can happen if a laboratory result is delayed, or a patient's electronic flowsheet is started in the middle of the treatment. We have added the ability to create new columns of data, and are designing the changes to the temporal processing components of ONCOCIN to allow for data that is inserted out of order. We have also extended the flowsheet to allow for patient specific parameters (e.g., special test results or symptoms) that the physician wishes to follow over time. The flowsheet layouts have been modified to create protocol specific flowsheets, e.g., lymphoma flowsheets have a different configuration than lung cancer flowsheets. The basic structure of the interface has been modified to use object-oriented methods, which allows for more flexible interaction between different components of the flowsheet and the operations performed on the flowsheet.

  A continuing area of research concerns how to guide the user to the most appropriate items to enter (based on the needs of the reasoning program) without disrupting the fixed layout of the flowsheet. The mainframe version of ONCOCIN modified the order of items on the flowsheet to extract necessary information from the user. In the workstation version, we have developed a guidance mechanism which alerts the user to items that are needed by the reasoning program. The user is not required to deviate from a preferred order of entry nor required to respond to a question for which no current answer is available. Cliff Wulfman is primarily responsible for improvements to the user interface of ONCOCIN.

- *System support for the reorganization.* The LISP language, which we used to build the first version of ONCOCIN, does not explicitly support basic knowledge manipulation techniques (such as message passing, inheritance techniques, or other object-oriented programming structures). These facilities are available in some commercial products, but none of the existing commercial implementations provide the reliability, speed, size, or special memory-manipulation techniques that are needed for our project. We have therefore developed a "minimal" object-oriented system to meet our specifications. The object system is currently in use by each component of the new version of ONCOCIN and in the software used to connect these components. In addition, all ONCOCIN student projects are now based on this programming environment. Christopher Lane created and is responsible for modifications to the object-oriented system.

C.2.2 Interactive Entry of Chemotherapy Protocols by Oncologists (OPAL)

A major effort in this grant year has been the continued development and testing of software (the OPAL system) that will permit physicians who are not computer programmers to enter protocol information on a structured set of forms presented on a graphics display. Most expert systems require tedious entry of the system's knowledge. In many other medical expert systems, each segment of knowledge is transferred from the physician to the programmer, who then enters the knowledge into the expert system. We have taken advantage of the generally well-structured nature of cancer treatment plans to design a knowledge entry program that can be used directly by clinicians. The structure of cancer treatment plans includes:

- choosing among multiple protocols (that may be related to each other);
- describing experimental research arms in each protocol;
- specifying individual drugs and drug combinations;
- setting the drug dosage level;
- and modifying either the choice of drugs or their dosage.

Using the graphics-oriented workstations, this information is presented to the user as computer-generated forms which appear on the screen. After the user fills in the blanks on the forms, the program generates the rules used to drive the reasoning process. As the user describes more detailed aspects of the protocol, new forms are added to the computer display; these allow the user to specify the special cases that make the protocols so complicated. Although the user is unaware of the creation of the knowledge base from the interaction with OPAL, a complex set of translations are taking place. The user's entries are mapped into an intermediate data structure (IDS) that is common for all protocols. From the IDS, a translation program generates rules for creating and modifying treatment, and integrates them with the existing ONCOCIN knowledge base. Improving the design of the IDS and the rule translation programs will be a major research effort of this year.

Although the "forms" were specifically designed for cancer treatment plans, the techniques used to organize data can be extended to other clinical trials, and eventually to other structured decision tasks. The key factor is to exploit the regularities in the structure of the task (e.g., this interface has an extensive notion of how chemotherapy regimens are constructed) rather than to try to build a knowledge-entry system that can accept *any* possible problem specification. The OPAL program is based upon a domain-independent forms creation package designed and implemented by David Combs. This program will provide the basis for our extension of OPAL to other application areas.

We have now entered thirty-five protocols covering many different organ systems and styles of protocol design (increased from 6 in last year's annual report). Based on this experience, we are modifying OPAL to increase the percentage of the protocol that can be entered directly by our clinical collaborators. One direction in which we have extended the OPAL program is in providing a graphical interface of nodes and arcs to specify the procedural knowledge about the order of treatments and important decision points within the treatments. This work is described in several papers by Musen.

C.2.3 Strategic Therapy Planning (ONYX)

As mentioned above, we have continued our research project (ONYX) to study the therapy-planning process and to determine how clinical strategies are used to plan therapy in unusual situations. Our goals for ONYX are: (1) to conduct basic research into the possible representations of the therapy-planning process, (2) to develop a computer program to represent this process, and (3) eventually to interface the planning program with ONCOCIN. We have worked with our clinical collaborators to determine how to create therapy plans for patients whose special clinical situation preclude following the standard therapeutic plan described in the protocol document.

The prototype program design has four components: (1) to review the patient's past record and recognize emerging problems, (2) to formulate a small number of revised therapy plans based on existing problems, (3) to determine the results of the generated plans by using simulation, and (4) to weight the results of the simulation and rank order the plans by performing decision analysis. This model is described in the papers by Langlotz.

We have built an expert system based on decision analytic techniques as part of the solution to the fourth step of the ONYX planning problem. The program carries out a dialogue with the user concerning the particular treatment choices to be compared, potential problems with the treatments, and the patient-specific utilities corresponding to the possible outcomes. A decision tree is automatically created, displayed on the screen, and solved. The solution is presented to the user, and is compatible with a explanation program for decision trees being developed as part of the Ph.D. research of Curtis Langlotz.

C.2.4 Documentation

In 1986, we videotaped a lecture and demonstration of the ONCOCIN and OPAL systems at the XEROX Palo Alto Research Center. This videotape is available for loan from our offices. Our previous videotapes have been shown at scientific meetings and have been distributed to many researchers in other countries. The publications described below further document our recent work on ONCOCIN.

C.2.5 Dissemination

We are planning experimental installation of ONCOCIN workstations in private oncology offices in San Jose and San Francisco. An application proposing this project is currently under review.

D. Publications Since January, 1986

1. Musen, M.A., Rohn, J.A., Fagan, L.M., and Shortliffe, E.H. Knowledge engineering for a clinical trial advice system: Uncovering errors in protocol specification (Memo KSL-85-51). Proceedings of AAMSI Congress 86 (A. Levy and B. Williams, eds.), pp. 24-27, Anaheim, 8-10 May 1986.

2. Langlotz, C.P., Fagan, L.M., and Shortliffe, E.H. Overcoming limitations of

artificial intelligence planning techniques. Memo KSL-85-52. Proceedings of AAMSI Congress 86 (A. Levy and B. Williams, eds.), pp. 92-96, Anaheim, 8-10 May 1986.

3. Musen, M.A., Fagan, L.M., and Shortliffe, E.H. Graphical specification of procedural knowledge for an expert system. Memo KSL-85-53. Presented at the Second IEEE Computer Society Workshop on Visual Languages, pp. 167-178, Dallas, TX, June 1986. Reprinted in Expert Systems: The User Interface (J. Hendler, ed.). Norwood, NJ: Ablex Publishing Company, 1987.

4. Langlotz, C.P., Fagan, L.M., Tu, S.W., Sikic, B.I., and Shortliffe, E.H. A therapy planning architecture that combines decision theory and artificial intelligence techniques. KSL-85-55. Submitted for publication, November 1986.

5. Combs, D.M., Musen, M.A., Fagan, L.M., and Shortliffe, E.H. Graphical entry of procedural and inferential knowledge. Memo KSL-85-56. Proceedings of AAMSI Congress 86 (A. Levy and B. Williams, eds.), pp. 298-302, Anaheim, 8-10 May 1986.

6. Lane, C.D., Frisse, M.E., Fagan, L.M., and Shortliffe, E.H. Object-oriented graphics in medical interface design. Memo KSL-85-58. Proceedings of AAMSI Congress 86 (A. Levy and B. Williams, eds.), pp. 293-297, Anaheim, 8-10 May 1986.

7. Musen, M.A., Fagan, L.M., Combs, D.M., and Shortliffe, E.H. Facilitating knowledge entry for an oncology therapy advisor using a model of the application area. Memo KSL-86-1. Proceedings of MEDINFO-86, pp. 46-50, Washington, D.C., October 1986.

8. Langlotz, C.P., Fagan, L.M., Tu, S.W., Sikic, B.I., and Shortliffe, E.H. Combining artificial intelligence and decision analysis for automated therapy planning assistance. Memo KSL-86-3. Proceedings of MEDINFO-86, pp. 794-798, Washington, D.C., October 1986.

9. Kahn, M.G., Fagan, L.M., and Shortliffe, E.H. Context-specific interpretation of patient records for a therapy advice system. Memo KSL-86-4. Proceedings of MEDINFO-86, pp. 175-179, Washington, D.C., October 1986.

10. Musen, M.A., Fagan, L.M., Combs, D.M., and Shortliffe, E.H. Use of a domain model to drive an interactive knowledge-editing tool. Memo KSL-86-24. To appear in the International Journal of Man-Machine Studies, 1987.

11. Langlotz, C.P., Shortliffe, E.H., and Fagan, L.M. Using decision theory to justify heuristics. Memo KSL-86-26. Proceedings of AAAI-86, pp. 215-219, Philadelphia, August 1986.

12. Shortliffe, E.H. Artificial Intelligence in Management Decisions: ONCOCIN. Memo KSL-86-39. Proceedings of a Conference on Medical Information Sciences, University of Texas Health Sciences Center at San Antonio, July 1985. To appear in Frontiers of Medical Information Sciences, Praeger Publishing, 1986.

13. Lane, C. The Ozone (O_3) Reference Manual. KSL-86-40, July 1986.

14. Musen, M.A., Combs, D.M., Walton, J.D., Shortliffe, E.H., and Fagan, L.M. OPAL: Toward the computer-aided design of oncology advice systems. Memo KSL-86-49. Proceedings of the Tenth Annual Symposium on Computer Applications in Medical Care, pp. 43-52, Washington, D.C., October 1986. Reprinted in Topics in Medical Artificial Intelligence (P.L. Miller, ed.), New York: Springer-Verlag, 1987.

15. Shortliffe, E.H. Medical expert systems: Knowledge tools for physicians. Memo KSL-86-52. Special issue on Medical Informatics, West. J. Med. 145:830-839, 1986.

16. Shortliffe, E.H. Medical expert systems research at Stanford University. Memo KSL-86-53. Presented at the Twentieth IBM Computer Science Symposium, Shizuoka, Japan, October 1986.

17. Langlotz, C.P., Shortliffe, E.H., and Fagan, L.M. A methodology for computer-based explanation of decision analysis. Working paper, KSL-86-57, November 1986.

18. Shortliffe, E.H. Computers in support of clinical decision making. Memo KSL-87-25, 1986. To appear in Lippincott's forthcoming Textbook of Internal Medicine (W.N. Kelley, ed.).

19. Langlotz, C.P. and Shortliffe, E.H. The relationship between decision theory and default reasoning. Working paper KSL-87-17, 1987.

20. Shortliffe, E.H. Computer programs to support clinical decision making. Memo KSL-87-30. To appear in JAMA, July 1987.

E. Funding Support

Grant Title: "Therapy-planning strategies for consultation by computer"
Principal Investigator: Edward H. Shortliffe
Project Management: Lawrence M. Fagan
Agency: National Library of Medicine
ID Number: LM-04136
Term: April 1987 to March 1990
Total award: $380,123

Grant Title: "Knowledge Management for Clinical Trial Advice Systems"
Principal Investigator: Edward H. Shortliffe
Project Management: Lawrence M. Fagan
Agency: National Library of Medicine
ID Number: 1 R01 LM04420-01
Term: September 1985 through August 1988
Total award: $314,707

Grant Title: Postdoctoral Training in Medical Information Science
Principal Investigator: Edward H. Shortliffe
Project Management: Edward H. Shortliffe
Agency: National Library of Medicine
ID Number: 1 T32 LM07033
Term: July 1, 1984 - June 30, 1989
Total award: $903,718

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 1988
Total award: $250,000 ($50,000 annually).

Grant Title: Explanation of Computer-assisted therapy plans
Principal Investigator: Lawrence M. Fagan
Agency: National Institutes of Health
ID Number: 1 R23 LM04316
Term: 2/1985-1/1988
Total award: $107,441

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. We also demonstrated our developing workstation code in the Xerox exhibit in the trade show associated with AAAI-84 in Austin, Texas, IJCAI-85 in Los Angeles, AAAI-86 in Philadelphia, and Medinfo 86. 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*, and ONCOCIN has been featured on the "NBC Nightly News," the PBS "Health Notes" series, and "The MacNeil-Lehrer Report." Most recently it appeared in a special on Artificial Intelligence for TV Ontario (Canadian PBS station). Due to the frequent requests for ONCOCIN demonstrations, we have produced a videotape about the ONCOCIN research which includes demonstrations of our professional workstation research projects and the 2020-based clinic system. The tape has been shown at several national meetings, including the 1984 Workshop on Artificial Intelligence in Medicine, the 1984 meeting of the Society for Medical Decision Making, and the 1985 meeting of the Society for Research and Education in Primary Care Internal Medicine. The tape has also been shown to both national and international researchers in biomedical computing. We have also completed an updated tape.

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 has been demonstrated via networks at both national and international meetings in the past, and several medical school and computer science teachers continue to use the program in their computer science or medical computing courses. Researchers who visit our laboratory often begin their introduction by experimenting with the MYCIN/EMYCIN systems. 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 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 which 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 dissemination 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 researchers on opposite coasts of the country.

During this past two years, we have had extensive interactions with Randy Miller at Pittsburgh. Via floppy disks and SUMEX, we have experimented with several versions of the QMR program. The interaction was very much facilitated by the availability of SUMEX for communication and data transmission.

C. Critique of Resource Management

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 diligently 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 Tom Rindfleisch for maintaining such high professional standards. As our computing needs grow, we have increased our dependence on special SUMEX skills such as networking and communication protocols.

III. RESEARCH PLANS

A. Project Goals and Plans

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

1. *Development of a workstation model for cost-effective dissemination of clinical consultation systems.* To meet this specific aim we will continue the basic and applied programming efforts (ONCOCIN, OPAL, and ONYX) described earlier in this report.
2. *To encode and implement for use by ONCOCIN the commonly used chemotherapy protocols from our oncology clinic.* In the upcoming year, we will:
    - Extend the OPAL protocol entry system
    - Continue entry of additional protocols at the rate of one protocol/month (including testing)
3. *To continue testing of the workstation version of ONCOCIN.*
4. *To generalize the reasoning and interaction components of the ONCOCIN system for other applications.*

B. Justification and Requirements for Continued SUMEX Use

All the work we are doing (ONCOCIN plus continued use of the original MYCIN program) continues to be dependent on daily use of the SUMEX resource. Although much of the ONCOCIN work has shifted 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.

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).

The main development of our project will continue to take place on LISP machines which we have purchased or which have been donated by the XEROX Corporation.

C. Requirements for Additional Computing Resources

The acquisition of the DEC 2020 by SUMEX was crucial to the growth of our research work. It ensured 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 ("D-machines") produced by Xerox, and our work will increasingly transfer to them over the next several years. Our current funding supports 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 on upgraded workstations for extremely CPU-intensive tasks. Although our dependence on SUMEX for workstations has decreased due to a recent gift from XEROX, our requirements for network support of the machines has drastically increased. Individual machines do not provide sufficient space to store all of the software used in our project, nor to provide backup or long-term storage of work in progress. It is the networks, file storage devices, protocol converters, and other parts of the SUMEX network that hold our project together. In addition, with a research group of about 20 people, we are taking advantage of file sharing, electronic mail, and other information coordinating activities provided by the DEC 2060. We hope that with systems support and research by SUMEX staff, we will be able to gradually move away from a need for the central coordinating machine over the next five years.

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 central machine and to the research workstations on which we are currently building the new ONCOCIN prototype. The D-machine's greater 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 workstations have provided a reliable, constant "load-average" machine for running experiments with physicians and for development work. The development of ONCOCIN on the D-machine 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 D-machine use wherever possible, and providing us with systems-level support when needed. We look forward to the arrival of additional advanced workstations and the development of a more distributed computing environment through SUMEX-AIM.

E. Responses to Questions Regarding Resource Future

1. "What do you think the role of the SUMEX-AIM resource should be for the period after 7/86, e.g., continue like it is, discontinue support of the central machine, act as a communications crossroads, develop software for user community workstations, etc.?"

   We believe that the trend towards distributed computing that characterized the early 1980's will continue during the second half of the decade. Although we have begun this process by moving much of our research activity to LISP machines, the SUMEX DEC-20 continues to be a major source of support for all communication, collaboration, and administrative functions. It also continues to provide a quality LISP environment for rapid prototyping, student projects in the early stages before workstations are made available, and for demonstrating system features to people at a distance. These latter functions are still not well handled by distributed machines, and we believe that a logical role for the resource in the future is to develop software and communications techniques that will allow us to further decrease our dependence on the large central machine.

2. "Will you require continued access to the SUMEX-AIM 2060 and if so, for how long?"

   As indicated above, our needs could still be met with a gradual phaseout of the 2060 over the next 3-5 years, provided that current services such as file handling and backup, mail, document preparation, and advanced network support are available from other machines (e.g., SAFE file server plus the Medical Computer Science file server). This implies maintenance of an ARPANET connection, connections to other campus machines, and facilities for linking together the heterogeneous collection of computing equipment upon which our research group depends. SUMEX would need to concentrate on providing software support for networks and systems software for workstations if it were to provide the same level of service we now experience while moving to a fully distributed environment.

3. "What would be the effect of imposing fees for using SUMEX resources (computing and communications) if NIH were to require this?"

   Since all our research is NIH-supported, we see nothing but administrative headaches without benefits if there were to be a move to require fee-for-service billing for access to shared SUMEX resources. The net effect would simply be a transfer of funds from one arm of NIH to another (assuming that the agencies that currently fund our work could supplement our grants to cover SUMEX charges), and there would be a simultaneous restraining effect on the research environment. The current scheme permits experimentation and flexibility in use that would be severely inhibited if all access incurred an incremental charge.

4. "Do you have plans to move your work to another machine workstation and if so, when and to what kind of system?"

As mentioned above, and described in greater detail in our annual report, we are making a major effort to move much of our research activity to LISP machines (currently Xerox 1108's, 1186's and HP-9836's). Our familiarity with this technology, and our commitment to it, have resulted solely from the foresight of the SUMEX resource in anticipating the technology and providing for it at the time of their last renewal. However, for the reasons mentioned above, we continue to depend upon the central communication node for many aspects of our activities and could effectively adapt to its demise only if the phaseout were gradual and accompanied by improved support for a totally distributed computing environment.

IV.A.4. PROTEAN Project

PROTEAN Project

Oleg Jardetzky
Nuclear Magnetic Resonance Lab, School of Medicine
Stanford University

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

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goals of this project are related both to biochemistry and artificial intelligence: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallography proteins), and (b) use protein structure determination as a test problem for experiments with the AI problem solving structure known as the Blackboard Model. Empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure. This problem exhibits considerable complexity, yet there is reason to believe that AI programs can be written that reason much as experts do to resolve these difficulties [12].

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 allow the study of proteins whose structure cannot be determined with other techniques, and will decrease the time needed for the determination.

C. Highlights of Progress

During the past year, we have expanded our initial prototype program, called PROTEAN, designed on the blackboard model. It is implemented in BB1 (discussed in the Core AI Research section of this report), a framework system for building blackboard systems that control their own problem-solving behavior.

The reasoning component of PROTEAN directs the actions of the Geometry System (GS), a set of programs that performs the computationally intensive task of positioning portions of a molecule with respect to each other in three dimensions. The GS runs in the UNIX environment on a Silicon Graphics IRIS 3020 graphics workstation, which provides computing performance comparable to a VAX 11/780 for our task. The reasoning program (in Lisp in BB1) is coupled to the GS by a local area computer network, maintained by SUMEX.

Pictures of the results of GS computations are displayed on the graphics screen of the IRIS workstation, using a locally developed program called DISPLAY to draw the evolving protein structures at several levels of detail. The DISPLAY program can be used to view structures generated by the GS either under the direct control of the user or as directed by the reasoning system running in BB1. MIDAS and MMS are two other molecular modeling and display systems to manipulate protein structures, particularly those obtained from crystallographic techniques as found in the Protein Data Bank. The ability to observe structures in three dimensions is essential to understanding the behavior of the PROTEAN's reasoning and geometry systems and provides essential insights on the problem solving process.

PROTEAN embodies the following experimental techniques for coping with the complexities of constraint satisfaction:

1. The problem-solver partitions each problem into a network of loosely-coupled sub-problems. PROTEAN first positions individual pieces of structures and their immediate neighbors within local coordinate systems. It subsequently composes the most constrained partial solutions developed for these sub-problems in a complete solution for the entire protein. This partitioning and composition technique reduces the combinatorics of search.

2. The problem-solver attempts to solve sub-problems and coordinate solutions at multiple levels of abstraction. For example, PROTEAN operates at two levels of abstraction. At the "Solid" level, it positions elements of the protein's secondary structure: alpha-helices, beta-sheets, and coils. At the "Atom" level, it positions the protein's individual atoms. Partial solutions at the solid level reduce the combinatorics of search at the lower level. Conversely, tightly constrained partial solutions at the lower level introduce new constraints on solid level solutions.

3. The problem-solver preserves the "family" of solutions consistent with all constraints applied thus far. For example, in positioning a helix within a partial solution, PROTEAN does not attempt to identify a unique spatial position for the helix. Instead, it identifies the entire spatial volume within which the helix might lie, given the constraints applied thus far. Preserving the family of legal solutions accommodates problems with incomplete constraints; the solution is constrained only as the data indicate. It also accommodates incompatible constraints by permitting disjunctive sub-families, which may be necessary for flexible proteins.

4. The problem-solver applies constraints one at a time, successively restricting the family of solutions hypothesized for different sub-problems. PROTEAN successively applies constraints on the positions of protein structures, restricting spatial volumes within which they may lie. This allows the different kinds of constraints to be applied by integrating their effects on a family of solutions.

5. The problem-solver tolerates overlapping solutions for different sub-problems. For example, in identifying the volume within which structure-a might lie in partial solution 1, PROTEAN may include part of the volume identified for structure-b. Overlapping volumes for two structures indicate either: (a) that the two structures actually occupy disjoint sub-volumes that cannot be distinguished within the larger, overlapping volumes identified for them because the constraints are incomplete; or (b) that the two structures are mobile and alternately occupy the shared volume.

6. The problem-solver reasons explicitly about control of its own problem-solving actions: which sub-problems it will attack, which partial solutions it will expand, and which constraints it will apply. Control reasoning guides the problem-solver to perform actions that minimize computation, while maximizing progress toward a complete solution. It also provides a foundation for the problem-solver's explanation of problem-solving

activities and intermediate partial solutions and for its learning of new control heuristics.

Multiple blackboards in PROTEAN allow several sets of knowledge to be used. A biochemical knowledge base stores information about proteins and secondary structures, amino acids, and atoms. A concept blackboard describes a concept hierarchy of natural types, object types, role types, contexts, constraint types, and problem solving methods. The ACCORD language blackboard explicitly represents the actions that can be taken in the language for arrangement assembly problems. The problem blackboard describes the protein to be solved and all experimental data observed for the molecule. Finally, the evolving solution of the protein structure is built on a third solution blackboard.

PROTEAN determines the structure of a protein by assembling the protein from components at several levels of detail. Initially, the major secondary structures of the protein are positioned relative to each other by considering them as solid structures, ignoring the side chains of the amino acids and representing constraints with respect to atoms of the protein backbone. This *solid level* approximation is sufficient to determine the overall shape of the molecule, but leaves details of the structure indistinct. Second, an *atomic level* representation of the protein including side chains is used with more precise distance, bond length, and bond angle constraints to remove chemically infeasible structures generated at the solid level. The atomic level description allows a more detailed description of the structure, at the cost of larger numbers of components to consider and increased computation time.

The reasoning component of PROTEAN includes domain and control knowledge sources for the assembly of a protein. Each domain knowledge source directs a small portion of the construction of the molecule. These knowledge sources develop partial solutions that position alpha helices, beta strands, and coils at the *solid level* and refine the resulting state families using all available distance constraints. Control knowledge sources determine which of the possible assembly actions is the best to perform at each stage of the problem solving.

We have built a first extension to PROTEAN that assembles a protein at the level of the atomic backbone. The facilities available include programs to manipulate protein data bank files and generate test data automatically, use atomic level constraints to prune solid level solutions, generate example instances of the protein backbone from the solid level structures, and generate candidate structures for unstructured coil segments of a protein. Work is in progress to combine the atomic level of assembly with the solid level to provide additional constraints at the more abstract level of assembly.

The PROTEAN system has been used to construct a complete solution at the solid level of detail for the Lac-repressor headpiece, a protein with fifty-one amino acids consisting of four coil sections and three alpha helices. In this work, the constraints were determined experimentally from NMR studies.

In addition to the Lac-repressor headpiece protein, we have applied PROTEAN to sperm whale myoglobin, T4 lysozyme, and cytochrome B. Each of these latter proteins has a known crystal structure. In each case, we extracted features of the protein structure and distance constraints from the crystal structure to build data sets for PROTEAN. We then applied the PROTEAN system to the resulting data sets to determine the behavior of the system with different kinds of input.

To determine the correctness and capabilities of the PROTEAN method, we have applied PROTEAN to sperm whale myoglobin, a molecule whose crystal structure is known. In this test, we used distance constraints that would be measured as NOEs, overall size information, and the interaction between the heme group and the amino acids. We also systematically explored the dependence of the precision and accuracy of the solutions on the quality of the input data available. In all cases, the solutions obtained from PROTEAN enclose the actual structure of the molecule, with the best results coming from data that includes many short range constraints.

We have also defined representations for structures such as the heme group in myoglobin and other cofactors that can be used in constraint satisfaction operations to further restrict the positions of the secondary structures in the protein.

The PROTEAN system takes the secondary structure as input. For molecules in solution, the extent of the helical, sheet, and unstructured coil segments of a protein is derived largely from NMR data between backbone and side chain hydrogen atoms. We have developed a knowledge-based system called ABC that uses heuristic knowledge and NMR data to automate this important step in protein structure determination. ABC is implemented using the BB1 blackboard architecture. In addition to solving the secondary structure classification problem, ABC provides a flexible and extensible framework for experimenting with identification methods for secondary structures as well as for data interpretation and pattern recognition techniques.

Work is proceeding on several aspects of the protein structure problem, including assembly of several partial arrangements and integration of these pieces of solution into larger structures, using atomic level volume exclusion of atoms and information on sidechain packing to produce more precise atomic level solutions, and developing more appropriate representations for unstructured coil sections of proteins.

D. Relevant Publications

1. Altman, R. and Jardetzky, O.: *New strategies for the determination of macromolecular structures in solution.* Journal of Biochemistry (Tokyo), Vol. 100, No. 6, p. 1403-1423, 1986.

2. Altman, R. and Buchanan, B.G.: *Partial Compilation of Control Knowledge.* To appear in Proceedings of the AAAI 1987.

3. Brinkley, J., Cornelius, C., Altman, R., Hayes-Roth, B. Lichtarge, O., Duncan, B., Buchanan, B.G., Jardetzky, O.: *Application of Constraint Satisfaction Techniques to the Determination of Protein Tertiary Structure.* Report KSL-86-28, Department of Computer Science, 1986.

4. Brinkley, James F., Buchanan, Bruce G., Altman, Russ B., Duncan, Bruce S., Cornelius, Craig W.: *A Heuristic Refinement Method for Spatial Constraint Satisfaction Problems.* Report KSL 87-05, Department of Computer Science.

5. Buchanan, B.G., Hayes-Roth, B., Lichtarge, O., Altman, A., Brinkley, J., Hewett, M., Cornelius, C., Duncan, B., Jardetzky, O.:*The Heuristic Refinement Method for Deriving Solution Structures of Proteins.* Report KSL-85-41. October 1985.

6. Garvey, Alan, Cornelius, Craig, and Hayes-Roth, Barbara: *Computational Costs versus Benefits of Control Reasoning.* Report KSL 87-11, Department of Computer Science.

7. Hayes-Roth, B.: *The Blackboard Architecture: A General Framework for Problem Solving?* Report HPP-83-30, Department of Computer Science, Stanford University, 1983.

8. Hayes-Roth, B.: *BB1: An Environment for Building Blackboard Systems that Control, Explain, and Learn about their own Behavior.* Report HPP-84-16, Department of Computer Science, Stanford University, 1984.

9. Hayes-Roth, B.: *A Blackboard Architecture for Control.* Artificial Intelligence 26:251-321, 1985.

10. Hayes-Roth, B. and Hewett, M.: *Learning Control Heuristics in BB1.* Report HPP-85-2, Department of Computer Science, 1985.

11. Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewett, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O.: *PROTEAN: Deriving protein structure from constraints.* Proceedings of the AAAI, 1986, p. 904-909.

12. Jardetzky, O.: *A Method for the Definition of the Solution Structure of Proteins from NMR and Other Physical Measurements: The LAC-Repressor Headpiece.* Proceedings of the International Conference on the Frontiers of Biochemistry and Molecular Biology, Alma Alta, June 17-24, 1984, October, 1984.

13. Lichtarge, Olivier: *Structure determination of proteins in solution by NMR.* Ph.D. Thesis, Stanford University, November, 1986.

14. Lichtarge, Olivier, Cornelius, Craig W., Buchanan, Bruce G., Jardetzky, Oleg: *Validation of the First Step of the Heuristic Refinement Method for the Derivation of Solution Structures of Proteins from NMR Data.*, April 1987. Submitted to Proteins: Structure, Function, and Genetics.

E. Funding Support

Title: Interpretation of NMR Data from Proteins Using AI Methods

PI's: Oleg Jardetzky and Bruce G. Buchanan

Agency: National Science Foundation

Grant identification number: DMB-8402348

Total Award Period and Amount: 2/1/87 - 9/30/89 $120,000
(includes direct and indirect costs)

Current award period and amount: 2/1/87 - 9/30/89 $120,000
(includes direct and indirect costs)

The following grants and contracts each provide partial funding for PROTEAN personnel.

Title: Modeling Exper Control

PI: Bruce G. Buchanan

Agency: Office of Naval Research

Grant Identification Number: ONR N00014-86-K-0652

Total award period and amount: 6/1/85 - 5/31/85, $96,879
(direct and indirect)

Current award period and amount: 6/1/85 - 5/31/85, $96,879
(direct and indirect)
PROTEAN component is $48,440 (direct & indirect) or 50% of grant

Title: Research on Blackboard Problem-Solving Systems

PI's: Edward A. Feigenbaum and Bruce G. Buchanan

Agency: Boeing Computer Services Corporation

Grant identification number: W-271799

Total award period and amount: 8/1/86 - 7/31/87, $245,432
(direct and indirect)

Current award period and amount: 8/1/86 - 7/31/87, $245,432
(direct and indirect)
PROTEAN component is $12,730 (direct & indirect) or 5% of grant

Title: Knowledge-Based Systems Research

PI: Edward A. Feigenbaum

Agency: Defense Advanced Projects Research Agency

Grant identification number: N00039-86-0033

Total award period and amount: 10/1/85 - 9/30/88 $4,130,230 (in negotiation)
(direct and indirect)

Current award period and amount: 10/1/86 - 9/30/87 $1,549,539
(direct and indirect)
PROTEAN component is $29031, or 1.9 % of grant total

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

We are occasionally in contact with researchers at Robert Langridge's laboratory at the University of San Francisco.

C. Critique of Resource Management

The SUMEX staff has continued to be most cooperative in supporting PROTEAN research. The SUMEX computer facility is well maintained and managed for effective support of our work. The computer network and Lisp workstations are supported very effectively by the SUMEX staff.

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 of NMR data on small proteins. The current version of PROTEAN has domain and control knowledge sources that implement the reasoning techniques described above to build a solution using a dynamically created strategic plan. These knowledge sources develop partial solutions that position multiple alpha helices, coils, and beta structures at the Solid level and refine those helices using distance, surface, and volume constraints.

PROTEAN also includes programs that use atomic level representations of the amino acid backbone and side chains. These routines use more precise atomic level distance constraints to prune the solutions obtained by the more abstract solid level geometry computations. Programs are also available to find acceptable backbone segments for unstructured coil segments between alpha helices and beta structures.

The proposed research would expand PROTEAN to include knowledge sources that:

1. merge highly constrained partial solutions at the Solid level.
2. propagate emergent constraints at the atomic level back up to the solid level to further restrict the relative positions of superordinate helices, beta sheets, and coils.
3. further restrict the relative locations of atoms relative to one another.
4. select instances of structures to be used as starting points for other kinds of refinement procedures, such as the solution of the Bloch equations, which define the NMR spectrum that can possibly arise from a given structure. These equations provide a very strong test of the correctness of our method, as well as providing an additional constraint on proposed structures.
5. develop efficient and effective control strategies for the solution of intermediate and large molecules.
6. reason about mobility of structures when the data indicate that mobility is possible.

We have built an effective strategy for automatically determining the families of solid level solutions for small proteins, such as the Lac-repressor headpiece. We will extend the current work to develop control strategies to guide PROTEAN's constraint satisfaction in medium and large protein to identify the family of legal protein conformations as efficiently as possible.

B. Justification for continued SUMEX use

We will continue to use SUMEX for developing parts of the program before integrating them with the whole system. We are using Interlisp to implement PROTEAN within the Blackboard model flexibly and quickly. In addition, the local area network that SUMEX maintains is crucial to the communications between our reasoning system in BB1, running on Xerox Lisp machines, and our geometry programs and display systems, running on the IRIS 3020 workstation.

C. Need for other computing resources

At this time our computational resources are almost adequate. However, access to Lisp machines for program development is often a limiting factor in our ability to continue the research. In addition, faster computation of the operations of the GS would be facilitated by a special-purpose array processor or an additional workstation for computing.

IV.A.5. 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 objectives of the RADIX project are 1) Discovery: to provide knowledgeable assistance to a research investigator in studying medical hypotheses on large databases, and to automate the process of hypothesis generation and exploratory confirmation, 2) Summarization: to develop a program and set of techniques for automated summarization of patient records, and 3) Peer Review: to develop a program to assist physician reviewers examine case databases for medical peer review and quality assurance. For system development we have used a subset of the ARAMIS database. We will first describe our work on discovery, followed by summarization and peer review.

RADIX Discovery Module

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 and many other clinical database projects 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.

More recently, we have undertaken a new component to the RADIX program: a knowledge-based discovery module. The goal of the knowledge-based discovery module is to overcome some of the limitations of the original, statistics-based, RX discovery module. In creating disease hypotheses, researchers make extensive use of notions of causation, mechanism of action, tempo, and quantitative sufficiency, as well as detailed knowledge of pathophysiology. We are seeking to automate this process of hypothesis formation by replicating selected discoveries in rheumatology using data from the ARAMIS database.

RADIX Summarization Module

The management of inpatients and outpatients is often complicated by the size and disorganization of patient charts. The current paper chart is ill-suited to serve as the major means of communication among health care providers. In recognition of this problem, computerized patient records are becoming increasingly available. While computerization of records at least renders them legible and available, it does not solve the problem of information overload. The ability to automatically create patient summaries would represent a useful adjunct to a patient record for rapid review of a case, for clinical decision making and patient monitoring, and for surveillance of quality of care. The goal of the RADIX summarization program is to infer a summary of a patient's clinical history from lengthy on-line medical records.

The RADIX summarization program is a knowledge-based system which produces intelligent summaries from a time-oriented data base of Systemic Lupus Erythematosus patients. Medical concepts in the system are represented by three entities of increasing complexity: abnormal primary attributes, abnormal states and diseases. Abnormal states and diseases are derived from the abnormal primary attributes by the Reasoner using a combination of model-driven and data-driven algorithms. Uncertainty associated with the derived states is handled with a Bayesian approach supplemented by boolean predicates, using likelihood ratios obtained from a transformation of the INTERNIST knowledge base. After summarizing the data, the system generates interactive, graphical displays with optional explanation windows.

The prototypes we have implemented have shown that intelligent summarization of medical records is feasible and that interactive graphical display is of great help in conveying complex medical information. However, the system is still under development and has not been formally evaluated. There is much work remaining to be done in the process of creating a complete, clinically useful summary. The knowledge base must be tested and enlarged, the temporal aspect of the reasoning must be improved and more sophisticated displays must be developed. Finally, although our program currently works only with the ARAMIS data base, we hope to extend it and produce a General Summarization System that could be interfaced with any time-oriented medical data base. This general system would include other data base dictionaries and would allow the user to enter medical knowledge tailored to his data base.

RADIX Peer Review Program

We have begun design of a program to assist physician reviewers with medical peer review and quality assurance. This work builds on the Summarization module, and extends it with a new Screening module. The Summarization module, described above, will allow a reviewer to rapidly scan a detailed, longitudinal record. It will summarize major events in the record by displaying them as labels on a time line. The new Screening module will take as input a reviewer's specification of rules of practice that he is interested in checking in the records. The module will transform these rules into an internal form in which they will be matched against the patient records. The output will be a set of episodes in the patient record in which apparent violations of the rules of practice have occurred. The reviewer will then be able to interactively examine each of these episodes using the Summarization module to determine whether a violation was substantiated by the context in which the medical decision was made.

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, and so forth.

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 in terms 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.

The medical relevance of the automated summarization program is readily apparent. A practicing physician or medical researcher, faced with a patient chart, often with dozens of visits and scores of attributes, rarely has time to read the entire chart. He (or she) would like a succinct summary of the important events in that patient's record to assist his decision making. The use of computerized medical records improves the quality of information but does not solve the problem of information overload. For this reason, it would be useful to have the ability to automatically summarize patient records into meaningful clinical events.

C. Highlights of Research Progress

C.1 April 1986 to April 1987

Our primary accomplishments in this period have been the following:

1) Design and implementation of a second generation of the automated summarization program.

2) Design and implementation of a bit-mapped display program for chronic patient data.

3) Development of algorithms for transforming the Internist knowledge base into standard Bayes forma.

4) Design of a Peer Review program based on the Summarization program.

5) Publication of papers on automated discovery and automated summarization, and presentation of results at medical conferences.

6) Training post-doctoral researchers, participants in RADIX, in methods of medical artificial intelligence research.

C.1.1 Design and implementation of a second generation of the prototype automated summarization program

We have designed and implemented a second generation of our prototype automated summarization program. This work is described in Dezegher-Geets, 1987, noted in the publications section. The current program improves upon a prototype implemented by Downs (Downs 1986); the knowledge base has been substantially enlarged, the inference mechanisms refined and enhanced for temporal reasoning, and the graphical display capability has been expanded. The summarization program produces intelligent summaries from a time-oriented data base of Systemic Lupus Erythematosus patients. Medical concepts in the system are represented by three entities of increasing complexity: abnormal primary attributes, abnormal states and diseases. Abnormal states and diseases are derived from the abnormal primary attributes by the Reasoner using a combination of model-driven and data-driven algorithms. Uncertainty associated with the derived states is handled with a Bayesian approach supplemented by boolean predicates, using likelihood ratios obtained from a transformation of the INTERNIST knowledge base. After summarizing the data, the system generates interactive, graphical displays with optional explanation windows.

C.1.2 Design and implementation of a bit-mapped display program for chronic patient data

The new display program provides graphic, synoptic, intelligent displays of chronic patient data. The goals of our implementation are:

1) Provide a good approximation of what each user actually wants and needs to see, without excess data.

2) Provide "intelligent" grouping of attributes based on knowledge of groups of related attributes, for example related to organ system, differential diagnoses, manifestations, and evidence.

3) Provide "intelligent" selection of attributes by prioritizing and selecting attributes by their clinical importance for the patient.

4) Provide interactive, editable displays, with choices available immediately through menus for the common displays.

The architecture is designed so that the Display Module sits "on top" of the AI components. It is designed to interact with a separate knowledge base or "expert system". The Display is separated from the knowledge base specifically to make it transportable and generalizable.

The knowledge based component contains knowledge of diseases, disease hierarchies, causal relations, equivalence relationships (e.g. proteinuria is part of Nephrotic syndrome), and so on. The display module has information that such relationships exist in medicine, and when to request specific information from the knowledge base. The Display module's knowledge of general medical concepts that are relevant for display includes the severity, belief, import, differential of a manifestation, complications of a disease, manifestations, organ system or user-specified attribute groupings, causal relationships, and equivalence relationships.

C.1.3 Development of algorithms for transforming the Internist knowledge base into standard Bayes form

INTERNIST-1 is an expert system for diagnosis across a broad spectrum of disease. Over twenty man-years of effort have gone into the construction of its knowledge base which contains relationships between approximately 600 diseases and 4,000 manifestations of disease. A major limitation of INTERNIST-1 is that the quantities used within the system to represent uncertainty, called evoking strengths and frequencies, are poorly defined. This makes it difficult to tune the method used by the program to assign likelihoods to diseases (the scoring scheme) and makes it difficult to transport knowledge contained in the program to other medical diagnostic systems. We have carried out several experiments involving the use of probability theory to better characterize the quantities. These experiments have been performed in collaboration with R. Miller and D. Heckerman.

In one experiment, assessments of p(D), p(M|D), and p(M|not D) were provided for approximately 100 manifestation-disease pairs representative of the knowledge base by Dr. Randy Miller, one of the principal contributors to the INTERNIST-1 project. These assessments were used to calculate positive likelihood ratios L(D|M), negative likelihood ratios L(D|not M), and posterior odds O(D|M). Using these assessments and calculated quantities, a graphical method was used to show that the evoking strength is more closely related to the likelihood ratio L(D|M) than to the posterior odds O(D|M).

In another experiment, a Chi-squared analysis showed that monotonic transformations of the evoking strength into positive likelihood ratios are significantly better than transformations into posterior odds, confirming the results of the previous experiment. It was also determined that monotonic transformations of frequency into p(M|D) are better than transformations into negative likelihood ratios.

Most recently, we attempted to optimize the transformation of the numbers in the INTERNIST KB into a probabilistic form. Various combinations of multiple regressions were performed on the evoking strengths, frequencies, probabilities of disease, p(D), and probabilities of manifestation, p(M), versus the likelihood ratios L(D|M) and L(D|not M). This process yielded some interesting and unexpected results. For example, the multiple regression of evoking strength AND p(M) vs. L(D|M) showed an r-squared of .84, significantly better than the r-squared value for evoking strength vs. L(D|M) alone. Also, the transformation from frequency, p(M), and p(D) into L(D|not M) revealed a correlation coefficient of .58. These results suggest a low cost method for converting the knowledge in INTERNIST-1 to a probabilistic form. In particular, assessments of p(D) and p(M) (only about 4500 numbers) can be used in conjunction with evoking strengths and frequencies in the KB (about 40,000 numbers) to construct likelihood ratios. We are currently testing a subset of the knowledge base to determine whether or such a conversion will improve the diagnostic performance of INTERNIST-1.

C.1.4 Design of a Peer Review program based on the Summarization program

We have begun design of a program to assist physician reviewers with medical peer review and quality assurance. This work builds on the Summarization module, and extends it with a new Screening module. The Summarization module, described above, will allow a reviewer to rapidly scan a detailed, longitudinal record. It will summarize major events in the record by displaying them as labels on a time line. The new Screening module will take as input a reviewer's specification of rules of practice that he is interested in checking in the records. The module will transform these rules into an internal form in which they will be matched against the patient records. The output will be a set of episodes in the patient record in which apparent violations of the rules of practice have occurred. The reviewer will then be able to interactively examine each of these episodes using the Summarization module to determine whether a violation was substantiated by the context in which the medical decision was made.

C.1.5 Publication of papers on automated discovery and automated summarization, and presentation of results at medical conferences

In addition to the publications noted above, we have submitted and/or had accepted additional papers, noted in the section on publications, and presented results at numerous medical conferences.

C.1.6 Training Post-Doctoral researchers, participants in RADIX, in methods of medical artificial intelligence research

We have been training three post-doctoral researchers on the project during the current reporting year; Andrew G. Freeman, M.D., Isabelle de Zegher-Geets, M.D., and Donald Rucker, M.D.. Andrew Freeman has been responsible for the new Display program, and for developing the Internist transformation algorithms. Isabelle de Zegher-Geets will complete a thesis this June on Automated Summarization as part of Stanford's Medical Information Sciences program; Don Rucker will undertake a thesis in the coming year on the Peer Review program.

C.2 Research in Progress

Our current research carries forward the work in automated summarization and automated discovery described above. Specifically, we are 1) implementing the intelligent discovery module, and evaluating and modifying its design as we get initial results, and 2) substantially expanding the prototype automated summarization module to be able to deal with a full patient record. We continue to work on problems involved in the representation of medical knowledge, as part of developing the programs for summarization and discovery. These programs act both as test beds for the extant knowledge representation techniques, and forcing functions for the development of new techniques.

D. Publications

1. Blum, R. L., and Walker, M.G.: *LISP as an Environment for Software Design: Powerful and Perspicuous.* In Proceedings of the Tenth Annual Symposium on Computer Applications in Medical Care, pages 326-331. IEEE Computer Society, IEEE Service Center, Piscataway, New Jersey, October, 1986. Invited paper and lecture.

2. Blum, R. L., and Walker, M.G.: *Automated Medical Discovery from Clinical Databases: An Overview of the RADIX Project.* In Proceedings of the Fifth Toyobo Biotechnology Foundation Symposium: Artificial Intelligence in Medicine. The Toyobo Foundation, Tokyo, Japan, August, 1986. Invitational address.

3. Blum, R. L., and Walker, M. G.: *Automated Medical Discovery from Clinical Databases: an Overview of the RADIX Project.* In Proceedings of the First International Conference on Artificial Intelligence and Its Impacts in Biology and Medicine, pages 59-83. Groupement Scientifique pour le Developpement de l'Intelligence Artificielle en Languedoc-Roussillon, Montpellier, France, September, 1986. Invitational address.

4. Blum, Robert L. and Gio C.M. Wiederhold: *Studying Hypotheses on a Time-Oriented Clinical Database: An Overview of the RX Project.* In J.A. Reggia and S. Thurim: 'Computer Assisted Medical Decision-Making'; Springer Verlag, 1985, pp.245-253.

5. Blum, R.L.: *Two Stage Regression: Application to a Time-Oriented Clinical Database.* Knowledge Systems Laboratory Technical Report. 1985.

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

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

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

9. 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.

10. 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.

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

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

13. 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.

14. 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.

15. 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.

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

17. 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)

18. 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.

19. 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.

20. Downs, S., Walker, M.G., and Blum, R.L.: Automated Summarization of On-line Medical Records. In Proceedings of the Fifth World Congress on Medical Informatics (Medinfo), pages 800-804. Elsevier Science Publishers, 1986.

21. Downs, S.M.: *A Program for Automated Summarization of On-line Medical Records*, Master's thesis, Stanford University, 1986.

22. DeZegher-Geets, I.M., Freeman, A.G., Walker, M.G., Blum, R.L., and Wiederhold, G.C.M.: *Computer-aided Summarization of a Time-oriented Medical Data Base*. In Proceedings of the Third Annual Conference on Computerization of Medical Records. Institute for Medical Record Economics, 1987.

23. DeZegher-Geets, I.: *Intelligent Summarization of a Time-Oriented Medical Data Base*. Master's thesis, Stanford University, 1987. In preparation.

24. Freeman, A.G., Walker, M.G., and Blum, R.L.: *Techniques for Bitmap Displays of Chronic Patient Data* 1987. In preparation.

25. Freeman, A.G., Heckerman, D., and Miller, R.: *Transformation of the Internist Knowledge Base to Bayes Form* 1987. In preparation.

26. Kuhn, I., Wiederhold, G., Rodnick, J.E., Ramsey-Klee, D.M., Benett, S., Beck, D.D.: *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.

27. Shortliffe, E. H., Wiederhold, G., and Fagan, L.: *An Introduction To Medical Computer Science*. Addison-Wesley, 1987. In preparation.

28. Walker, M.G.: *How Feasible is Automated Discovery?* IEEE Expert 2(1):70-82, Spring, 1987.

29. Walker, M.G., and Blum, R.L.: *Towards Automated Discovery from Clinical Databases: the RADIX Project*. In Proceedings of the Fifth World Congress on Medical Informatics (Medinfo), pages 32-36. Elsevier Science Publishers, 1986.

30. Walker, M.G., Blum, R.L., and Fagan, L.M.: *Minimycin: A Miniature Rule-Based System*. M.D.Computing, Vol. 2, No. 4., 1985.

31. Walker, M.G., and Blum, R.L.: *An Introduction to LISP*. M.D. Computing, Vol. 2, No. 1., 1985.

32. Wiederhold, G.: *File Organization for Database Design*. McGraw-Hill Book Company, 1987.

33. Wiederhold, G.C.M., Walker, M.G., Hasan, W., Chaudhuri, S., Swami, A., Cha, S.K., Qian, X., Winslett, M., DeMichiel, L., and Rathmann, P.K. *KSYS: An Architecture for Integrating Databases and Knowledge Bases*. 1987. Submitted to IEEE Transactions on Software Engineering, special issue on Artificial Intelligence.

34. Wiederhold, G.C.M., Walker, M. G., Blum, R.L., Downs, S.M. and Dezegher-Geets, I.M.: *Acquisition of Medical Knowledge from Medical Records*. In Proceedings of Benutzergruppenseminar Medizin Systems 87. Benutzergruppenseminar Medizin, Munich, Germany, October, 1987.

35. Wiederhold, G.: *Views, Objects, and Databases*. IEEE Computer 19(12):37-44, December, 1986.

36. Wiederhold, G.C.M., Walker, M. G., Blum, R.L., and Downs, S.M.: *Acquisition of Knowledge from Data*. In Ras, Z.W., and Zemankova, M. (editors), Proceedings of the International Symposium on Methodologies for Intelligent Systems. University of Tennessee, Knoxville, Tennessee, October, 1986.

37. Wiederhold, Gio, Robert L. Blum, and Michael Walker: *An Integration of Knowledge and Data Representation.* Proc. of Islamorada Workshop, Feb. 1985, Computer Corporation of America, Cambridge MA; Chapter 23 of 'On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies' (Brodie, Mylopoulos, and Schmidt, eds.), Springer Verlag, June 1986.

38. Wiederhold, Gio: *Knowledge versus Data.* Chapter 9 of 'On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies' (Brodie, Mylopoulos, and Schmidt, eds.) Springer Verlag, Feb. 1986.

39. Missikoff, Michele and Gio Wiederhold: *Towards a Unified Approach for Expert and Database Systems.* In 'Expert Database Systems', Larry Kerschberg (editor), Benjamin/Cummings, 1986, pages 383-399; also in Proceedings of First Workshop on Expert Database Systems, Kiawah Island, South Carolina, Oct. 1984, vol.1, pp.186-206.

40. Wiederhold, Gio and Paul D. Clayton: *Processing Biological Data in Real Time.* M.D. Computing, Springer Verlag, Vol.2 No.6, November 1985, pages 16-25.

41. Wiederhold, Gio: *Knowledge Bases.* Future Generations Computer Systems, North-Holland, vol.1 no.4; April 1985, pp.223--235.

42. Wiederhold, Gio: *Disease Registers: Use of Databases to Generate New Medical Knowledge.* In K.Abt, W.Giere and B.Leiber: 'Krankendaten, Krankheitsregister, Datenschutz', vol.58, Medizinische Informatik und Statistik, Springer Verlag, 1985, pp.39-55.

43. Wiederhold, G.: *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.

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

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

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

E. Funding Support Status

1) Knowledge-Based Management Structures
Gio C. M. Wiederhold, Ph.D.: Principal Investigator
Department of the Navy: N00039-84-C-0211
Total award: $1,756,410
Term: 1987 through 1990

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

Once the RADIX programs are 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

During the current reporting year we have had frequent interaction with members of other SUMEX projects; for example, development of algorithms for transforming INTERNIST data to Bayes form, presentation of research results at Stanford Medical Information Science Colloquia, discussions of automated discovery and automated summarization, practical programming issues, and training of Medical Computer Science Students in the use of KEE, Lisp workstations, and so on. The SUMEX community is an invaluable resource for providing such interaction.

C. Critique of Resource Management

The DECSystem 20 continues to provide acceptable performance, but it is frequently heavily loaded at peak hours.

The SUMEX resource management continues to be accessible and quite helpful.

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 --

Our short-term goals focus on the two activities described earlier: implementation and further development of the intelligent discovery module, and substantial expansion of the automated summarization program to deal with an entire rheumatology patient record.

LONG-RANGE GOALS --

The long-range goals of the RADIX Project are 1) automatic discovery of knowledge in a large time-oriented database, and provision of assistance to a clinician who is interested in testing a specific hypothesis, and 2) development of techniques for automated summarization of patient records. We hope to make these programs sufficiently robust that they will work over a broad range of hypotheses and over a broad spectrum of 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.

We also would hope that the central SUMEX facility, the DEC 2060, would continue to be supported. We continue to make constant use of this machine for text-editing, document preparation, file and database handling, communications, and program demos.

Responses to Questions Regarding Resource Future

Q: What do you think the role of the SUMEX-AIM resource should be for the period after 7/87, e.g., continue like it is, discontinue support of the central machine, act as a communications crossroads, develop software for user community workstations, etc.

A: In our opinion, the SUMEX 2060 should continue to be supported. The machine continues to be of value to us for text-editing (TVedit and EMACS), for document preparation (SCRIBE), and for communications and mail. We also depend on it as a central, reliable facility for program demos, for manipulating large databases, and maintaining central program files. It would be a real loss if it was discontinued.

Software for community work stations. Yes. Making good utility programs available to all users sounds like a good idea.

Q: Will you require continued access to the SUMEX-AIM 2060 and if so, for how long?

A: Yes. For the foreseeable future and for the above reasons.

Q: What would be the effect of imposing fees for using SUMEX resources (computing and communications) if NIH were to require this?

A: We would pay them. The 2060 is worth it to us. Of course, if the fees were high, we would consider alternatives.

Q: Do you have plans to move your work to another machine workstation and if so, when and to what kind of system?

A: We are currently using two of the SUMEX Xerox 1108's for the development of our project. We will stay with these for the foreseeable future.

IV.B. 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.

IV.B.1. INTERNIST-I Project

CADUCEUS Project (INTERNIST-I)

This project is unfunded at the present time.

J. D. Myers, M.D.
University Professor Emeritus (Medicine)
University of Pittsburgh
1291 Scaife Hall
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-1, 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 highlighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The development of the QUICK MEDICAL REFERENCE (QMR) under the leadership of Dr. Randolph A. Miller has allowed us to distribute the INTERNIST-I knowledge base in a modified format to over twenty other academic medical institutions. The knowledge base can thereby be used as an "electronic textbook" in medical education at all levels -- by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The INTERNIST-I program has been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program, and to develop patient management problems and test cases for the Part III Examination and the developing computerized testing program of the National Board of Medical Examiners.

C. Highlights of Research Progress

C.1 Accomplishments this past year

For the record, it should be noted that grant support for the QMR project has come solely from the CAMDAT Foundation of Farmington, Conn., from the Department of Medicine of the University of Pittsburgh, and from Dr. Miller's NLM RCDA grant. The NLM and DRR grants currently supporting the CADUCEUS project do not in any way support the QMR project.

The group of us (Myers, Miller and Masarie) together with assigned residents in internal medicine and fellows in medical informatics are continuing to expand the knowledge base and to incorporate the diagnostic consultative program into QMR. The computer program for the interrogative part of the diagnostic program is the main remaining task. An editor for the QMR knowledge base, as modified from the INTERNIST-I knowledge base, has been written from scratch in Turbo Pascal by Dr. Masarie. The entire QMR program can be accommodated in, maintained (particularly edited) and operated on individual IBM PC-AT computers.

Our group has incorporated into the QMR diagnostic consultant program modifications and embellishments of the INTERNIST-I knowledge base, and will continue to do so over the next year by adding "facets" of diseases or syndromes. This addition and modification is expected to improve the performance of the diagnostic consultant program.

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. Several dozen new diseases have been profiled during the past year. The current number of diseases in the QMR knowledge base is 577, and over 4100 possible patient findings are included.

C.2 Research in progress

There are four 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 and the QMR modification.

2. Institution of field trials of QMR on the clinical services in internal medicine at the Health Center of the University of Pittsburgh. This has been accomplished in a limited fashion beginning April 1987; a "computer-based diagnostic consultation service" has been made available to attending physicians and housestaff on the medical services of our two main teaching hospitals. Institutional Review Board (IRB) approval was granted to the service before it was initiated.

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

4. Adaptation of the diagnostic program and data base of INTERNIST-I and the QMR modification 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, and preliminary evaluation of the diagnostic consultation service.

D. List of relevant publications

1. Myers, J.D.: Educating future physicians: Something old, Something new. Ohio State Univ. Proceedings of Symposium, Medical Education in the 21st Century. 1985.
2. Myers, J.D.: The process of clinical diagnosis and its adaptation to the computer. In The Logic of Discovery and Diagnosis in Medicine, University of Pittsburgh Series in the Philosophy and History of Science, edited by Kenneth F. Schaffner, Univ. of California Press, pp. 155-180, 1985.
3. Masarie, Jr. F.E., Miller, R.A., First, M.B., Myers, J.D.: An Electronic Textbook of Medicine. Proceedings of Ninth Annual Symposium on Computer Applications in Medical Care. Baltimore, Maryland, November 1985.
4. Masarie, Jr. F.E., Myers, J.D., Miller, R.A.: INTERNIST-I PROPERTIES: Representing Common Sense on Good Medical Practice in a Computerized Medical Knowledge Base. Computers and Biomedical Research. 18: 458-479, October 1985.
5. Myers, J.D., Chairman. Medical Education in the Information Age. Proceedings of the Symposium on Medical Informatics. Association of American Medical Colleges, 1986.
6. Miller, R.A., Schaffner K.F., Meisel, A. Ethical and Legal Issues Related to the Use of Computer Programs in Clinical Medicine. Annals of Internal Medicine. 1985; 102:529-36.
7. Miller, R.A., Masarie, F.E., Myers J.D. "Quick Medical Reference" for diagnostic assistance. MD Computing. 1986; 3:34-48.
8. Miller, R.A., Mc Neil, M.A., Challinor, S., Masarie, F.E., Myers, J.D. Status Report: The INTERNIST-1/Quick Medical Reference Project. West J Med. 1986; 145:816-22.
9. Miller, R.A. From Automated Medical Records to Expert System Knowledge Bases: Common Problems in Representing and Processing Patient Data. Topics in Health Record Management, March 1987, in press.
10. Masarie, F.E., Miller, R.A. Medical Subject Headings and Medical Terminology: An analysis of terminology used in hospital charts. Bulletin of the Medical Library Association, 1987; 75:89-94.
11. Miller, R.A., Masarie, F.E., Miller, R.A. Quick Medical Reference (QMR): A microcomputer-based adaptation of the INTERNIST-1 diagnostic system for general internal medicine. In: R. Salamon, B. Blum, M. Jorgenson (eds), MEDINFO 86, p. 1143. Amsterdam: North Holland Publishing Co, 1986.
12. Heckerman, D., Miller, R.A. Towards a better understanding of the INTERNIST-1 knowledge base. In: R. Salamon, B. Blum, M. Jorgenson (eds), MEDINFO 86, pp. 22-26. Amsterdam: North Holland Publishing Co, 1986.
13. McNeil, M.A., Challinor, S.M., Miller, R.A. Preliminary Evaluation of a computer-based medical decision support system in the clinical setting. In Press, Medical Decision-Making, December 1986.

E. Funding support

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

   5 R24 RR01101-08
   07/01/80 - 03/31/86 - $1,658,347
   07/01/84 - 09/30/85 - $354,211
   09/30/85 - 03/31/86 - $50,690

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

   5 R01 LM03710-05
   07/01/80 - 03/31/86 - $853,200
   07/01/84 - 09/30/85 - $210,091
   09/30/85 - 03/31/86 - $35,316

3. Diagnostic-Internist: A Computerized Medical Consultant
   Randolph A. Miller, M.D.
   Associate Professor of Medicine
   University of Pittsburgh Department of Medicine
   National Library of Medicine - Development Award Research
   Career
   National Institutes of Health

   1 KO4 LM00084-01
   09/30/85 - 09/29/90 - amounts to be determined annually
   09/30/85 - 09/29/86 - $55,296
   09/30/86 - 09/29/87 - $55,296

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A,B. Medical Collaborations and Program Dissemination Via SUMEX

INTERNIST-I and QMR remain in a stage of research and particularly 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 INTERNIST-I. 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 base in internal medicine 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 QMR. 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 adaptation of our programs to the IBM PC-AT. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended for the present to keep INTERNIST-1 at SUMEX for comparative use as QMR is developed here.

Our best prediction is that our project will require continued access to the 2060 for the next year or two and we consider such access essential to the future development of our knowledge base. After that time, our work can probably be accomplished on our personal work stations.

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

Our predictable needs in this area will be met by our newly acquired personal work stations.

IV.B.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 five years, Polson, in collaboration with Dr. David Kieras of the University of Michigan, has 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 developed formal representations of a user's knowledge of how to operate a device and of the user-device interface (Kieras & Polson, 1985) and have completed several experiments evaluating their theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986).

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 Michigan. 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, 1985).

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 van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in 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.

B. 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.

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, 1984; Polson, Muncher, and Engelbeck 1985). The results of this project should be directly relevant to the design of complex, computer controlled medical equipment. They 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 modeling is a stated goal of both SUMEX and research supported by NIMH.

C. Highlights of Research Progress

The version of the prose comprehension model of 1978 (Kintsch & van Dijk, 1978), which originally was realized as a computer simulation by Miller & Kintsch (1980), has been extended in a major simulation program by Young (1985). Unlike the earlier program, Young includes macroprocessing in her model, and thereby greatly extends the usefulness of the program. It is expected that this program will be widely useful in studies of prose where a detailed theoretical analysis is desired.

The general theory has been reformulated and expanded in van Dijk & Kintsch (1983). This research report of book length presents a general framework for a comprehensive theory of discourse processing. It has been applied to an interesting special case, the question of how children understand and solve word arithmetic problems, by Kintsch & Greeno (1985). A simulation for this model, using INTERLISP, has been supplied in Fletcher (1985).

The device complexity project is in its fifth year. They have developed an explicit model for the knowledge structures involved in the user-device interaction, and they are developing simulation programs. Their preliminary theoretical results are described in Kieras & Polson (1985). They have also completed several experiments evaluating the theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986) and have shown that number of productions predicts learning time and that number of cycles and working memory operations predicts execution time for a method.

D. List of Relevant Publications

1. Fletcher, R.C.: *Understanding and solving word arithmetic problems: A computer simulation.* Technical Report No. 135, Institute of Cognitive Science, Colorado, 1984.

2. Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity.* Int. J. Man-Machine Studies, 22, 365-394, 1985.

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, 1985, 92, 109-129.

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. 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.

7. 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, Instrumentation, & Computers, 1984, 16, 249-255.

8. Polson, P.G. and Kieras, D.E.: *A quantitative model of the learning and performance of text editing knowledge.* In Borman, L. and Curtis, B. (Eds.) Proceedings of the CHI 1985 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 207-212, 1985.

9. Polson, P.G. and Jeffries, R.: *Instruction in general problem solving skills: An analysis of four approaches.* In (Eds.) Siegel, J., Chipman, S., and Glaser, R. Thinking and learning skills: Relating instructions to basic research: Vol. 1. Hillsdale, N.J.: OpLawrence Erlbaum Associates, pp. 414-455.

10. Polson, P.G., Muncher, E., and Engelbeck, G.: *Test of a common elements theory of transfer.* In Mantei, M. and Orbeton, P. (Eds.) Proceedings of the CHI 1986 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 78-83, 1986.

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

12. Young, S.: *A theory and simulation of macrostructure.* Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

13. Walker, H.W., Kintsch, W.: *Automatic and strategic aspects of knowledge retrieval.* Cognitive Science, 1985, 9, 261-283.

E. Funding Support

1. Text Comprehension and Memory
   Walter Kintsch, Professor, University of Colorado
   National Institute of Mental Health - 5 R01 MH15872-14-16
   7/1/84 - 6/30/87: $197,500 (direct)

2. Understanding and solving word arithmetic problems
   Walter Kintsch, Professor, University of Colorado
   National Science Foundation
   8/1/83 - 7/31/86: $200,000
   8/1/86 - 7/31/87: $55,400

3. Theories, Methods, and Tools for the Design of User-centered Computer Systems
   Walter Kintsch, Professor, University of Colorado
   Gerhard Fischer, Assoc. Prof. University of Colorado
   Army Research Institute
   8/1/86 - 7/31/91: $500,000
   8/1/86 - 7/31/87: $86,500

4. Software Design for a Propositionalizer
   Walter Kintsch, Professor, University of Colorado
   A. Turner, Research Assoc., University of Colorado
   Air Force Office of Scientific Research
   10/1/85 - 9/30/87: $110,000
   10/1/86 - 9/30/87: $47,000

4. The Application of Cognitive Complexity Theory to the Design of User Interface Architectures
   David Kieras, Associate Professor, University of Michigan
   Peter G. Polson, Professor, University of Colorado
   International Business Machines Corporation
   1/1/85 - 4/31/87: $500,000 (direct+indirect)
   1/1/86 - 4/31/87: $250,000 (direct+indirect)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

B. 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.

C. Critique of Resource Management

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 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 van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in 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 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 user's 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

Both the prose comprehension and the user-computer interaction projects have shifted their actual simulation work from SUMEX to systems at the University of Colorado and the University of Michigan. Both projects use Xerox 1108 systems continuing their work in INTERLISP.

Access to SUMEX's mail facilities are critical for the continued success of these projects. These facilities provide us with the means to interact with colleagues at other universities. Kintsch is currently collaborating with James Greeno, who is at the University of California at Berkeley, and Polson's long-term collaborator, David Kieras, is at the University of Michigan. In addition, our access to the Xerox 1108 (Dandelion) user's community is through SUMEX.

We currently use five computing systems: a VAX 11/780, a MicroVAX II, and three Xerox 1108s, one of which is at the University of Michigan. The VAX's are used to collect experimental data designed to evaluate the simulation models and to do necessary statistical analysis.

C. Needs and Plans for Other Computational Resources

SUMEX provides us with communication which we discussed in the preceding paragraph.

D. Recommendations for Future Community and Resource Development

We will continue to need access to the SUMEX-AIM 2060 in order to access communication networks.

IV.B.3. MENTOR Project

MENTOR Project

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

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 record-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 a collaboration between Dr. Blaschke at Stanford University, Dr. Speedie at the University of Maryland, and Dr. Charles Friedman at the University of North Carolina. Dr. Speedie provides the expertise in the area of artificial intelligence programming. Dr. Blaschke provides the medical expertise. Dr. Friedman contributes expertise in the area of physician feedback design and system impact evaluation. 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 project has been funded by the National Center for Health Services Research since January 1, 1985. Initial effort focused on exploration of the problem of designing the MENTOR system. As of June 1, 1987, a working prototype system has been developed and is undergoing evaluation. The prototype consists of a Patient Data Base, an Inference Engine, an Advisory Module and a Medical Knowledge Base. The Medical Knowledge Base currently contains information related to Aminoglycoside Therapy, Digoxin therapy, Surgical Prophylaxis, and Microbiology Lab reports. The system is currently implemented on a Xerox 1186 AI Workstation. Another version of the Patient Data Base has been developed for a VAX 750 and is currently being tested. Plans call for the interconnection of the VAX and the 1186 running the inference engine. The VAX will then be connected to a Hospital Information System for data acquisition.

E. Funding Support

Title: MENTOR: Monitoring Drug Therapy for Hospitalized Patients

Principal Investigators:

Terrence F. Blaschke, M.D.
Division of Clinical Pharmacology
Department of Medicine
Stanford University

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

Funding Agency: National Center for Health Services Research

Grant Identification Number: 1 R18 HS05263

Total Award: January 1, 1985 - December 31, 1988 $485,134 Total Direct Costs

Current Period: January 1, 1987 - December 31, 1987 $195,731 Total Direct Costs

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, the University of Maryland School of Pharmacy, and the University of North Carolina in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, facilitates this collaboration of geographically separated project participants by providing electronic mail and file exchange between sites.

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 system development issues. Dr. Perry Miller has also been of assistance by providing software for advisory generation. 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 quite satisfied with the current resource management.

III. RESEARCH PLANS

A. Project Goals and Plans

The MENTOR 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. 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.

1987 will be spent on continued prototype development in four content areas, refinement of the inference mechanisms, and interfacing to existing patient information systems.

B. Justification and Requirements for Continued SUMEX Use

This project needs continued use of the SUMEX facilities for two reasons. First, it provides access to an environment specifically designed for the development of AI systems. The MENTOR project focuses on the development of such a system for drug monitoring that will explore some neglected aspects of AI in medicine. This environment is necessary for the timely development of a well-designed and efficient MENTOR system. Second, access to SUMEX is necessary to support the collaborative efforts of geographically separated development teams at Stanford and the University of Maryland.

Furthermore, the MENTOR project is predicated on the access to the SUMEX resource free of charge over the next two years. Given the current restrictions on funding, the scope of the project would have to be greatly reduced if there were charges for use of SUMEX.

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 have transported the prototype system to a dedicated hardware system that can fully support the the planned system and which can be integrated into a Hospital Information System. For this purpose a VAX 750 and three Xerox 1186 workstations have been acquired and our development efforts have been transferred to them.

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 clearly evident that the users' almost insatiable demands for CPU cycles and disk space cannot be met by a single central machine. The best strategy would appear to be one of emphasizing powerful workstations or relatively small, multi-user machines linked together in a nationwide network with SUMEX serving as the its central hub. This would give the individual users much more control over the resources available for their needs, yet at the same time allow for the communications among users that have been one of SUMEX's strong points.

For such a network to be successful, further work needs to be done in improving the network capabilities of SUMEX to encourage users at sites other than Stanford. Further work is also needed in the area of personal workstations to link them to such a network. Given the successful completion of this work, it would be reasonable to consider the gradual phase-out of the central SUMEX machine over two or three years and its replacement by an efficient, high-speed communications server.

IV.B.4. SOLVER Project

SOLVER: Problem Solving Expertise

Dr. P. E. Johnson
Center for Research in Human Learning
University of Minnesota

Dr. James R. Slagle
Department of Computer Science
University of Minnesota

Dr. W. B. Thompson
Department of Computer Science
University of Minnesota

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The SOLVER project is an interdisciplinary research effort concerned with understanding medical expertise, particularly in diagnostic tasks. 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 fifteen years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only a limited amount of expertise. 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 knowledge 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, financial auditing, management, and law.

B. Medical Relevance and Collaboration

Much of our research has been and will continue to be directly focused on medical AI problems. Medical diagnostic expertise is a complex phenomenon which is not yet fully understood. The SOLVER project is studying both the theoretical foundations of expertise and also is engaged in the design and testing of medical expert systems.

A medical expert system in pediatric cardiology has been designed in collaboration with Dr. James Moller, Department of Pediatrics, University of Minnesota Hospitals.

Dr. Donald Connelly, Department of Laboratory Medicine, University of Minnesota School of Medicine, has supervised a number of medical expert system projects, including projects in analysis of time series observations and platelet transfusion practice.

Dr. Slagle's research group has developed an expert system shell called AGNESS ("A Generalized Network-based Expert System Shell"), and has developed three medical expert systems that either use AGNESS or are modeled after AGNESS. AGNESS uses a computation network rather than a production rule base and supports values of any well-defined data type, the Merit questioning scheme, an explanation facility, and expert-defined inference methods. The first major application of AGNESS was to implement the clinical expert system ETA (Exercise Test Analyzer). The cases studied came from the Program on the Surgical Control of the Hyperlipidemas (POSCH), a study of the effect of reduced cholesterol on heart attack victims.

Kent Spackman, M.D. was a post-doctoral fellow in medical informatics at the University of Minnesota who is completing a Ph.D. thesis in Artificial Intelligence at the University of Illinois. During his residency at the University of Minnesota Hospitals, Dr. Spackman collaborated with the SOLVER project. Dr. Spackman's research addressed issues in automated knowledge acquisition for medical expert systems.

C. Highlights of Research Progress

Accomplishments of This Past Year

Dr. Connelly has continued supervising the development of an expert system, ESPRE, to be used in monitoring requests for platelet transfusions. The prototype knowledge base was refined and extended, communications protocols to communicate with laboratory computer systems have been improved, a standing order feature has been implemented, the inference engine has been modified, and a preliminary evaluation has been completed. In the evaluation, 68 transfusion requests were processed by the system. In more than 80% of the cases, the expert system agreed with the blood bank decision to transfuse platelets. In six of the remaining cases, the expert system declined to propose a decision because there was no recent platelet count available to it. In four cases, additional clinical factors known to the blood bank physician were brought to bear, and the transfusions were authorized even though usual transfusion criteria had not been met. The expert system is being placed in parallel operation in the blood bank to be used as a consultation tool.

In addition, Dr. Connelly is supervising the project dealing with detection of deviations in time series by the human observer. This project involves the implementation of a number of small expert systems used in modeling the human graph reader. During the past year the work has been extended by examining individual observer differences in deviation detection performance and approach to graph reading. Time trend graphs representing monthly monitoring of serum carcinoembryonic antigen (CEA) levels in simulated patients with surgically-removed breast cancer were presented to twelve clinical laboratory observers and a time series analysis (TSA) routine which is based on a homeostatic model. The observers described their rationale in assigning a level of suspicion regarding the presence of an important deviation as the observation points were serially revealed to them. The verbalization reports were analyzed to develop rules that described each reader's graph reading strategy. Strategies were compared for commonality and difference. Rules obtained from the top two observers were merged into a common rule base and an expert system implemented. A second expert system was constructed by merging consistent rules from all observers. The deviation detection performance of all three approaches (TSA, observers, and both expert systems) are being compared. The analysis is currently in progress.

Dr. Johnson's research group has developed an expert system inference engine called "Cleric." Cleric is a rule based language written in Common Lisp which resembles a forward chaining production system. Cleric has been written to investigate diagnostic problem solving tasks. Cleric differs from simple production systems because it can dynamically create new specialized forms of existing rules for later execution. In addition, Cleric uses a subset of de Kleer's assumption based truth maintenance system (ATMS). A computer hardware diagnosis expert system called "Vesalius" has been implemented in Cleric.

The ETA (Exercise Test Analyzer) expert system has been implemented and tested. The change in the health of the patient's heart, as measured by treadmill ECG tests, between any two tests was rated on a seven-point scale; each subject was rated on several features and overall. Rules for sub-area ratings were built from the verbal protocols of a POSCH cardiologist, and then weightings for combining sub-area ratings into an overall rating were determined. ETA was tested on 100 cases from the POSCH study and outperformed both the average POSCH cardiologist and a previously developed multiple regression model.

In the past year, the expert system ESCA ("Evaluator of Serial Coronary Angiograms") has been developed with domain knowledge organized in an inference network modeled after that of AGNESS. The domain knowledge was gathered from verbal protocols of a POSCH member inferring changes in atherosclerotic disease from changes in the flow of blood as revealed in angiograms taken at different times. In some cases, the POSCH member was first asked to determine the change solely from a form recording the consensus of a two-member sub-panel, and then was shown a more detailed and less stylized diagram and allowed to modify his conclusion. A sub-panel working from the films was also observed so the influence of the perceptual component could be judged. Indeed, much of ESCA's success is due to factoring the domain into a perceptual component followed by an expert system component. Its success thus dispels doubts about the applicability of expert system technology to domains with significant perceptual components. ESCA performed slightly better than the sub-panel of clinicians for the cases examined. Using ESCA for subjective clinical evaluation, and one cardiologist to screen the conclusions, POSCH can now evaluate films faster, more consistently, and with less cost.

Research in Progress

The research in progress for the current year will be a continuation of projects that have been underway for some time. The main areas will be --

1. *Inference engine mechanisms in diagnostic reasoning.* This will be a continuation of the Cleric/Vesalius project. The Cleric language will be used to model different diagnostic strategies -- path-following, compare and conquer, and stateless analysis.

2. *Merit system for question selection.* AGNESS is being used in developing an expert system for early detection of clinical trends in cystic fibrosis (CF) patients. In addition, the ESCA expert system will be extended to consider multiple lines of reasoning and to make use of the Dempster-Shafer method.

3. *Detection of deviations in time series by the human observer.* Surveillance and early detection of deviation from a homeostatic state are goals common to health care programs for the apparently healthy as well as for groups of patients known to have or have had specific diseases. Automated approaches to detecting deviations have the advantage of being reliably applied, traceable, consistent in outcome, and conserving of professional resources. Rule based expert systems based upon analysis of human graph reading strategies are being evaluated.

4. *Knowledge based system for improving transfusion practice.* The ESPRE expert system has undergone preliminary evaluation and is now being used in parallel with traditional decision processes in transfusion therapy.

D. List of Relevant Publications

1. Bailey, A. et al.: *Auditing, artificial intelligence, and expert systems*, DECISION SUPPORT SYSTEMS: THEORY AND APPLICATION, A.B. Whinston (ed.), North-Holland Publications, 1985.

2. Connelly, D.P. and Rich, S.S.: *Detection of deviations in time series by the human observer*. Proc. IMIA Workshop on Maintaining a Healthy State within the Individual. (in press)

3. Connelly, D. and Johnson, P.E.: *Medical problem solving*. Human Pathology, 11(5):412-419, 1980.

4. 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.

5. 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.

6. 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, Addison-Wesley, 1984, pp. 275-319.

7. Johnson, P.E. and Zualkernan, I.: *Building expert systems for troubleshooting from knowledge of the design process*, IASTED International Conference on Expert Systems, Geneva, Switzerland, June, 1987.

8. Johnson, P.E., Zualkernan, I., and Garber,S.: *Specification of expertise*, International Journal of Man-Machine Studies (in press).

9. Johnson, P.E. *Cognitive models of expertise*, Proceedings of USC Symposium on Expert Systems and Auditor Judgment, February, 1986.

10. Johnson, P.E.: *Problem Solving*. In ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY, McGraw-Hill, 1985.

11. Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis*. In Bolc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

12. 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.

13. Johnson, P.E.: *What kind of expert should a system be?* Journal of Medicine and Philosophy, 8:77-97, 1983.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. Johnson, P.E., Hassebrock, F. and Moller, J.H.: *Multimethod study of clinical judgement*. Organizational Behavior and Human Performance 30:201-230, 1982.

20. 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.

21. Sielaff, B.H., Galen, G., Scott, E. and Connelly, D.P.: *Knowledge-based system for improving transfusion practice*. In AAMSI Congress 87 (in press).

22. Simpson, D.E., Rich, E., Dalgaard, K., Gjerdingen, K., Crowson, T., O'Brien, D., Johnson, P.E.: *The diagnostic process in primary care: A comparison of general internists and family physicians*. SOCIAL SCIENCE AND MEDICINE (in press).

23. Slagle, J. et al.: "Expert systems in medical studies -- A new twist," *Proceedings of the Conference on Applications of Artificial Intelligence*, SPIE, 1986.

24. Slagle, J.: *An expert system for a resource allocation problem*, Communications of the ACM, September, 1985.

25. Slagle, J.: *Alpha Beta Pruning*, ENCYCLOPEDIA OF ARTIFICIAL INTELLIGENCE, S. Shapiro (ed.), John Wiley and Sons, Inc., New York, 1985.

26. Slagle, J. and Gaynor, M.: *An intelligent control strategy for computer consultation*, IEEE Trans. on Pattern Analysis and Machine Intelligence, March, 1984.

27. Spackman, K.A. and Connelly, D.P.: *Knowledge-based systems in laboratory medicine and pathology: A review and survey of the field*. Archives of Laboratory Medicine and Pathology 111:116-119, 1987.

28. 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.

29. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

E. Funding and Support

Work on the SOLVER project is currently supported by grants from the Control Data Corporation ($95,000; 1986-88) and IBM ($81,000; 1987) to Paul Johnson ($95,000; 1986-88) and by a grant from the Microelectronics and Information Sciences Center (MEIS) at the University of Minnesota to Paul Johnson, William Thompson, James Slagle ($300,000; 1986-7).

Research in medical informatics is supported, in part, by a training grant from the National Library of Medicine, LM-00160, in the amount of $712,573 for the period 1984-1989. Dr. Connelly and Prof. Johnson are participants in this grant. The post-doctoral fellowship of Dr. Spackman was funded by this grant.

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.

The Galen system is available on SUMEX from the University of Minnesota as an unsupported research tool for the study of recognition based reasoning systems.

B. Sharing and Interactions with Other SUMEX-AIM Projects

The SOLVER project has not been engaged in any formal sharing with other projects in the last year. The SUMEX resource has continued to serve as a communications vehicle for informal contacts with other researchers. Dr. Johnson conducted informal conferences during the year with Drs. Bruce Buchanan and William Clancey.

C. Critique of Resource Management

None.

III. RESEARCH PLANS

A. Project Goals and Plans

An overall goal of the project is to describe methods for the specification of expertise. Our objective is to construct an artifact (for example, an expert system) that can solve a class of problems which is currently solved by an expert. To construct this artifact a specification of the requirements is needed which outlines what needs to be computed to solve the problem.

A number of artifacts may achieve the same performance in a variety of ways. The expert's method works because it is adapted to the capabilities of the human information processing system and the demands of the problem-solving task. Since we may implement our specification on various kinds of processors, we seek a description that does not depend on a particular processing architecture. The purpose of knowledge acquisition is not to learn how to solve a problem, but rather to discover *what is required* to solve a problem.

Our goal is to use protocol records of problem-solving activity to develop a specification of the requirements for any artifact that would attempt to solve the same problem. Given a class of problems, such as medical diagnostic tasks, and a protocol record from experts solving these problems, the task is to determine a method for transforming the protocol into a specification of expertise.

Our goal is to investigate the following framework for specification of expertise:

1. The expert can be viewed as a processor that has the capability of producing certain problem-solving behavior using expertise. The task of knowledge acquisition is to determine this expertise.
2. The expert has developed a set of actions and abilities that are necessary to realize this expertise.
3. Although we cannot observe the expertise directly, we can observe the invocation of the expert's actions and abilities in a record of problem-solving behavior.
4. Since we can observe the invocation of actions and abilities by the expert, we can develop some representation of the expertise.
5. A statement of the expertise required to perform a task serves as a specification of the requirements for a computer program that is designed to perform the task.

The development of a specific methodology for collecting and analyzing protocol data to arrive at a formal specification of expertise.

B. Justification and Requirements for Continued SUMEX Use

Our current model development takes advantage of the sophisticated Lisp programming environments on SUMEX and local facilities. 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.

We have completed a CommonLisp implementation of the Galen system and expect to rely heavily on CommonLisp for future projects.

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

Our current research support has permitted us to purchase Sun workstations for our Artificial Intelligence laboratory. The availability of CommonLisp on these machines is one reason why we expect to make use of that language in the future.

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

As a remote site, we particularly appreciate the communications that the SUMEX facility provides our researchers with other members of the community. We, too, are moving toward a workstation-based development environment, but we hope that SUMEX will continue to serve as a focal point for the medical AI community. In addition to communication and sharing of programs, we are interested in development of CommonLisp based knowledge engineering tools. The continued existence of the SUMEX resource is very important to us.

IV.B.5. ATTENDING Project

ATTENDING Project--Expert Critiquing Systems

Perry L. Miller, M.D. Ph.D.
Department of Anesthesiology
Yale University School of Medicine
New Haven, CT 06510

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

Our project is exploring the "critiquing" approach to bringing computer-based advice to the practicing physician.

Critiquing is a different approach to the design of artificial intelligence based expert systems. Most medical expert systems attempt to simulate a physician's decision-making process. As a result, they have the clinical effect of trying to tell a physician what to do: how to practice medicine. In contrast, a critiquing system first asks the physician how he contemplates approaching his patient's care, and then critiques that plan. In the critique, the system discusses any risks or benefits of the proposed approach, and of any other approaches which might be preferred. It is anticipated that the critiquing approach may be particularly well suited for domains, like medicine, where decisions involve a great deal of *subjective* judgment.

To date, several prototype critiquing systems have been developed in different medical domains:

1. ATTENDING, the first system to implement the critiquing approach, critiques anesthetic management.
2. HT-ATTENDING critiques the pharmacologic management of essential hypertension.
3. VQ-ATTENDING critiques aspects of ventilator management.
4. PHEO-ATTENDING critiques the laboratory and radiologic workup of a patient for a suspected pheichromocytoma.
5. In addition, a domain-independent system, ESSENTIAL-ATTENDING, has been developed to facilitate the implementation of critiquing systems in other domains.

C. Highlights of Research Progress

Current projects include the following:

HT-ATTENDING The original prototype version of HT-ATTENDING has been converted to the ESSENTIAL-ATTENDING format, and updated to reflect current thinking in the field of hypertension management. A major priority is to subject this system to validation and clinical evaluation, and to explore how best to disseminate the system as a practical consultation tool.

DxCON: Critiquing Radiologic Workup DxCON extends the design developed in PHEO-ATTENDING to critique the radiologic workup of suspected obstructive jaundice. Workup is an area in which we will aggressively pursue the critiquing approach for two reasons. 1) Since many areas of workup are quite constrained, it may prove possible to develop and test complete systems in a reasonably short time-frame. 2) Since workup is expensive, and very wasteful of resources if performed improperly, a computer system which helps to optimize a physician's workup plans could have significant economic benefits. The present national emphasis on controlling health costs makes this project very topical. We are also using this domain to explore issues of knowledge acquisition and verification.

ICON: Critiquing Radiological Differential Diagnosis Most existing diagnostic computer systems produce a ranked differential diagnosis as their output. In this process, the rich structure of the knowledge that went into developing the diagnoses may be lost to the user. ICON explores a different approach to diagnostic advice in the domain of radiology. To use ICON, a radiologist describes a set of findings seen on chest x-ray, together with a proposed diagnosis. ICON then produces a detailed analysis of *why* the observed findings serve to support or to rule out the diagnosis. It may also suggest further findings that might help refine the diagnosis, again explaining why the findings are important.

D. Publications

1. Miller, P.L.: Expert Critiquing Systems: Practice-Based Medical Consultation by Computer. New York: Springer-Verlag, 1986.

2. Miller, P.L. (Ed.): Selected Topics in Medical Artificial Intelligence. New York: Springer-Verlag (in press).

3. Miller, P.L., Shaw, C., Rose, J.R., Swett, H.A.: Critiquing the process of radiologic differential diagnosis. Computer Methods and Programs in Biomedicine 22:12-25, 1986.

4. Miller, P.L.: The evaluation of artificial intelligence systems in medicine. Computer Methods and Programs in Biomedicine 22:5-11, 1986.

5. Rennels, G.D., Shortliffe, E.H., Miller, P.L.: Choice of explanation in medical management: A multi-attribute model of artificial intelligence approaches. Medical Decision Making 7:22-31, 1987.

6. Rennels, G.D., Miller, P.L.: Artificial intelligence research in anesthesia and intensive care. Anesthesiology (submitted).

7. Miller, P.L., Rennels, G.D.: Prose generation from expert systems: An applied computational linguistics approach (submitted).

8. Mars, N.J.I., Miller, P.L.: Knowledge acquisition and verification tools for medical expert systems. Medical Decision Making 7:6-11, 1987.

9. Miller, P.L., Blumenfrucht, S.J., Rose, J.R., Rothschild, M., Swett, H.A., Weltin, G., Mars, N.J.I.: HYDRA: A knowledge acquisition tool for expert systems which critique medical workup. Medical Decision Making 7:12-21, 1987.

10. Swett, H.A., Miller, P.L.: ICON: A computer-based approach to differential diagnosis in radiology. Radiology (in press).

11. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: A computational model of reasoning from the clinical literature. Computer Methods and Programs in Biomedicine (in press).

12. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: A structured representation of the clinical literature and its use in a medical management advice system. Bulletin du Cancer (in press).

13. Miller, P.L., Barwick, K.W., Morrow, J.S., Powsner, S.M., Riely, C.A.: Semantic relationships and medical bibliographic retrieval: A preliminary assessment (submitted).

14. Miller, P.L.: Exploring the critiquing approach: Clinical practice-based feedback by computer. Biomedical Measurement, Informatics and Control (submitted).

15. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: A computational model of reasoning from the clinical literature. The AI Magazine (accepted pending revision).

16. Miller, P.L.: Exploring the critiquing approach: Sophisticated practice-based feedback by computer. Proceedings of the Fifth World Conference on Medical Informatics MEDINFO-86, Washington, D.C., October 1986, pp 2-6.

17. Mars, N.J.I., Miller, P.L.: Tools for knowledge acquisition and verification in medicine. Proceedings of the Tenth Symposium on Computer Applications in Medical Care, Washington, D.C., October 1986, pp. 36-42.

18. Miller, P.L., Blumenfrucht, S.J., Rose, J.R., Rothschild, M., Weltin, G., Swett, H.A., Mars, N.J.I.: Expert system knowledge acquisition for domains of medical workup: An augmented transition network model. Proceedings of the Tenth Symposium on Computer Applications in Medical Care, Washington, D.C., October 1986, pp. 30-35.

19. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: Reasoning from the clinical literature: The Roundsman system. Proceedings of the Fifth World Conference on Medical Informatics MEDINFO-86, Washington, D.C., October 1986, pp. 771-775.

20. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: Updating an expert knowledge base as medical knowledge evolves: Examples from oncology management. Proceedings of the American Association of Medical Systems and Informatics Congress-87, San Francisco, May 1987, pp. 238-231.

21. Fisher, P.R., Miller, P.L., Swett, H.A.: A script-based representation of medical knowledge involving multiple perspectives. Proceedings of the American Association of Medical Systems and Informatics Congress-87, San Francisco, May 1987, pp. 233-237.

22. Miller, P.L.: Expert consultation systems in medicine: A complex and fascinating domain. Proceedings of the Annual Meeting of the IEEE (Electro-87), New York, April 1987, pp. 1/2:1-4 (invited paper).

23. Miller, P.L., Fisher, P.R.: Causal models in medical artificial intelligence. Proceedings of the Eleventh Symposium on Computer Applications in Medical Care, Washington, D.C., November 1987 (submitted).

24. Powsner, S.M., Barwick, K.W., Morrow, J.S., Riely, C.A., Miller, P.L.: Coding semantic relationships for medical bibliographic retrieval: A preliminary study. Proceedings of the Eleventh Symposium on Computer Applications in Medical Care, Washington, D.C., November 1987 (submitted).

E. Funding Support

EXPERT COMPUTER SYSTEMS WHICH CRITIQUE PHYSICIAN PLANS
NIH Grant R01 LM04336
Principal Investigator: Perry L. Miller, M.D., Ph.D.
Annual Direct Costs: approximately $100,000
Period of Support: 9/1/85-8/31/87

This two-year grant supports the exploration of the critiquing approach to bringing computer-based advice to the physician, focusing primarily on the underlying system design issues.

SUPPORT OF THE UNIFIED MEDICAL LANGUAGE PROGRAM
NLM Contract N01-LM-6-3524
Principal Investigator: Perry L. Miller, M.D., Ph.D.
Annual Direct Costs: approximately $100,000
Period of Support: 8/22/86-8/21/88

This two-year research contract is part of the NLM Unified Medical Language (UML) program. We are defining a set of semantic relationships which could be used to augment the UML, to facilitate such functions as medical bibliographic retrieval.

SUPPORT FOR MEDICAL INFORMATICS AND ARTIFICIAL INTELLIGENCE
Ira DeCamp Foundation
Co-Principal Investigators: Henry A. Swett, M.D.
Perry L. Miller, M.D., Ph.D.
Annual Costs: $75,000
Period of Support: 7/1/86-6/30/90

This grant supports our present Medical Informatics program and is currently being used primarily to support Medical Informatics research training. If the present training application is funded, the Ira DeCamp support could be used for other activities in support of the training such as for a program secretary and for computing programming support.

MEDICAL INFORMATICS RESEARCH TRAINING AT YALE
Principal Investigator: Perry L. Miller, M.D., Ph.D.

We have been informed that we will receive a five-year training grant starting July 1, 1987.

Pending Support

EXPERT COMPUTER SYSTEMS WHICH CRITIQUE PHYSICIAN PLANS
Principal Investigator: Perry L. Miller, M.D., Ph.D.
Annual Direct Costs: approximately $100,000
Period of Support: 9/1/87-8/31/90

This grant requests continuation of our currently funded grant which is exploring the critiquing approach to bringing computer-based advice to the practicing physician. This continuation grant application focuses especially on refining and evaluating the HT-ATTENDING system which critiques hypertension management.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

Until recently we have been using the RUTGERS-AIM Resource. We used that facility to implement all of our early critiquing systems. We are currently in the early stages of moving part of our critiquing research to the SUMEX-AIM facility. Our main uses of SUMEX-AIM will be the following:

1. We will use SUMEX-AIM to demonstrate two of our systems, ATTENDING and HT-ATTENDING.
2. We will use SUMEX-AIM for the continued refinement of HT-ATTENDING, and for a planned controlled clinical experiment to measure the effect of HT-ATTENDING's advice on patient care. This will be performed in the Yale New Haven Hospital Primary Care Center, and is planned to commence this coming year.
3. We will use SUMEX-AIM for communication access to the national AIM community.

We have found our use of the RUTGERS-AIM facility to be extremely valuable. It provided us the resources needed to initiate our research and to continue several projects which are still active. It provided a natural vehicle to allow us to demonstrate the various systems easily, both in the United States and in Europe. Also, it enabled us to collaborate very closely with Dr. Glenn Rennels in his Stanford Medical Information Science thesis project on the Roundsman system. Via SUMEX-AIM and RUTGERS-AIM, Dr. Rennels and Dr. Miller maintained very close contact, typically with multiple messages each week, and sometimes within a single day.

III. FUTURE PLANS

We plan to continue our critiquing research as outlined above. One of our highest priorities will be the controlled experimental evaluation of the HT-ATTENDING system, which will be done using SUMEX-AIM. We will also continue to utilize SUMEX-AIM as outlined above. Although we are increasingly moving a great deal of our work onto internal workstations, we nevertheless plan to continue our use of SUMEX-AIM, especially in the further refinement and evaluation of HT-ATTENDING.

IV.C. 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.

IV.C.1. REFEREE Project

REFEREE Project

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

Byron W. Brown, Ph.D., Co-Principal Investigator
Department of Medicine
Stanford University

Daniel E. Feldman, Ph.D., M.D., Associate Investigator
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goals of this project are related both to medical science and artificial intelligence: (a) use AI methods to allow the informed but non-expert reader of the medical literature to evaluate a randomized clinical trial, and (b) use the interpretation of the medical literature as a test problem for studies of knowledge acquisition and fusion of information from disparate sources. REFEREE and REVIEWER, a planned extension, will be used to evaluate the medical literature of clinical trials to determine the quality of a clinical trial, make judgements on the efficacy of the treatment proposed, and synthesize rules of clinical practice. The research is an initial step toward a more general goal - building computer systems to help the clinician and medical scientist read the medical literature more critically and more rapidly.

B. Medical Relevance

The explosive growth of the medical literature has created a severe information gap for the busy clinician. Most physicians can afford neither the time required to study all the pertinent journal articles in their field, nor the risk of ignoring potentially significant discoveries. The majority of clinicians, in fact, have little sophistication in epidemiology and statistics; they must nonetheless base their pragmatic decisions on a combination of clinical experience and published literature. The clinician's computerized assistant must ferret out useful maxims of clinical practice from the medical literature, pass judgment on the quality of medical reports, evaluate the efficacy of proposed treatments, and adjudicate the interpretation of conflicting and even contradictory studies.

C. Highlights of Progress

REFEREE, a rule-based system built upon the EMYCIN framework, partially encodes the epidemiological knowledge of two highly regarded experts at Stanford, a biostatistician (Dr. Bill Brown) and a clinician (Dr. Dan Feldman). The REFEREE system, in particular, allows the informed but non-expert reader of the medical literature to evaluate the believability of a randomized clinical trial.

In the future, REFEREE and its extensions will alleviate the knowledge-acquisition bottleneck for an automated medical decision-maker: the program will evaluate the quality of a clinical trial, judge the efficacy of the treatment proposed therein, and synthesize rules of clinical practice. For the present, however, the fusion of knowledge from disparate sources remains a problem in pure AI. The efforts of the REFEREE team have instead focused their efforts on the refinement and deepening of REFEREE's biostatistical knowledge by applying effective knowledge acquisition and knowledge engineering techniques. Dr. Diana Forsythe and Dr. Harold Lehman are developing and using interview methods to acquire this knowledge from Dr. Brown, and R. Martin Chavez is implementing this in the prototype REFEREE expert system.

The REFEREE prototype is a consultant that evaluates the design and reporting of a single conclusion from randomized control trial for its believability. It contains, in preliminary form, Professor Brown's expert knowledge of biostatistics. REFEREE evaluates each statistical procedure described by the authors of the paper. The automated consultant then determines the most appropriate method for the problem at hand, based on the design of the trial and the hypotheses to be tested. REFEREE checks critical assumptions, looks for possible statistical abuses, verifies adjustments, and re-computes the statistics. In a beta-blocker study that employs the Cox proportional-hazards model, for instance, REFEREE will analyze the Kaplan-Meier survival curve and verify or reject the presence of a significant treatment effect.

The Knowledge Base: In order to evaluate the paper's presentation of a statistical test, REFEREE must apply three kinds of knowledge:

1. the statistical techniques that are relevant to the kinds of data likely to be found in a randomized clinical trial.
2. the methods to perform statistical tests to verify the paper's results.
3. the techniques to test hypotheses, to determine if the data in a paper support the conclusions of that paper.

Randomized controlled trials are used to test hypotheses regarding the effectiveness of various kinds of medical interventions. Dr. Brown classifies studies on the basis of three major attributes: the type of intervention tested (e.g. drug, surgery, health process change, etc.); the type of endpoint against which that intervention was tested (e.g. mortality, objective morbidity, subjective morbidity, etc.); and the type of conclusion drawn by the investigator/author on the basis of the research (e.g. that different treatments do or do not produce different outcomes, that a particular treatment is or is not cost-effective, etc.). Following this classificatory scheme, we decided to begin by producing a prototype REFEREE system that would help the reader to evaluate a single published conclusion concerning the effect of a given drug treatment on mortality.

Knowledge Acquisition: Having defined the scope of the initial knowledge base, we turned to the problem of collecting the information from Dr. Brown for inclusion in the system, i.e. knowledge acquisition. This task generally involves a relatively long-term process of face-to-face information gathering during sessions between the expert and one or more knowledge engineers. Dr. Diana Forsythe has noted a parallel between the communicative and analytical tasks involved in knowledge acquisition and those undertaken in ethnographic research. For this reason, we included an anthropologist in the research team and make use of ethnographic techniques in order to maximize the efficiency and quality of the data collection process.

Dr. Lehmann and Dr. Forsythe have carried out several months of systematic interviews with Dr. Brown in order to begin the process of constructing and refining the knowledge base for the current REFEREE prototype. We have combined a case-based approach that allows us actively to observe Dr. Brown as he reads papers, with semidirected interviewing oriented toward understanding his terminology and category system. We find that these techniques work very well: Dr. Brown's interest in the knowledge acquisition process has been sustained, and indeed has increased over time as the system based on his expertise has evolved. He is clearly comfortable with this approach, and notes that it has actually afforded him additional insight into the way he interprets the literature.

In order to codify the information gathered from Dr. Brown, Dr. Lehmann chose a model based on the *influence diagrams* used in decision analysis, in which the expert indicates which factors or *parameters* he finds crucial in making his judgement about the quality of the paper. Based on information from our expert, we have taken "believability" as the primary parameter of the present system, defined operationally by Dr. Brown as "the odds I am willing to give that the conclusions of the paper would be replicated in an experiment based on the methods reported in the paper but without any of the flaws". Within the influence diagram, parameters are connected to each other in a structure indicating the information considered by Dr. Brown in making particular judgments. In assessing believability, for instance, he considers the acceptability of the randomization, the quality of the blinding, other sources of bias, and how well the results substantiate the conclusion. Our use of influence diagrams has numerous advantages: the approach is acceptable to Dr. Brown, it is flexible, it can represent several aspects of the structure of the knowledge used by the expert, and the resultant data can be entered easily into the computer.

Once entered into the machine, the influence diagram is converted into *rules* such as the following:

```
If   : The quality of the randomization is high and
       The quality of the blinding is poor and
       The other sources of bias are unknown and
       The results substantiate the conclusion,

Then : There is suggestive evidence (0.7) that the believability of the
       clinical trial is high.
```

The number (0.7) captures the uncertainty of the expert in drawing a specific conclusion from the specific antecedents; this number is known as a *certainty factor*. The mathematics of certainty factors has been widely discussed in the literature.

Inference in REFEREE: REFEREE was originally built within EMYCIN, an AI environment developed from MYCIN at Stanford. In 1986 Chavez introduced some fundamental improvements to the REFEREE program; among other things, these changes greatly improved communication with the user (see "The User Interface", below).

The system is programmed to act as a problem solver, following the rules in the knowledge base in a *backwards chaining* path. For instance, the machine has the determination of the paper's believability as its *goal*. At the outset it finds a rule that reasons about the paper's believability (the above example). It then examines each antecedent of that rule in turn and looks for rules that draw a conclusion on *that* parameter, recursively, until an antecedent is found that has no rules. REFEREE then queries the user about that antecedent. For instance, from the rule "If the method of randomization was reported and the design of the randomization was good and the implementation of the randomization was poor - Then there is suggestive evidence (.6) that quality of the randomization method was acceptable", the machine would find that there are no rules that conclude that the method of randomization was reported. It would then ask the user, "Was the method of randomization reported?" If the answer is "No", then the machine abandons the rule in question, but saves the response for possible use with other rules. Note how this differs from a traditional paper-and-pencil checklist, for instance, where the user is confronted with each question regardless of its relevance.

The User Interface: The first versions of REFEREE were written to be used with a terminal connected to a large mainframe computer. In the past year Chavez has transformed the program so as to function at a stand-alone workstation. His first new version was written in an commercial expert system shell (KEE) which rested on an INTERLISP base; however, we then re-wrote the program for the Texas Instrument Explorer in CommonLisp.

The program code is now entirely independent of the knowledge required for reading papers. REFEREE has a new interface that is intuitive and consistent. There is an innovative consultation mode in which questions are presented in free-format menus. The dialogues are mixed-initiative and of mixed levels, allowing the user such options as requesting more detailed questions or cutting off apparently fruitless lines of questioning. With the new REFEREE prototype, the user interacts with the machine using a mouse-pointing device, as with the Macintosh. All questions are asked in a similar format. Finally, the screen enables the user to orient himself at all times, obviating the need for special commands to help the user "navigate" through the knowledge base. Our expert recently provided the best indication of the useability of this new system. After only a brief introduction to the new machine and interface, he was able - for the first time - to run an entire consultation by himself.

Current Status: At this point, REFEREE is a stable prototype that enables the clinician to read clinical trials more critically. As such, REFEREE represents only the first step in a larger research plan, the automation of knowledge acquisition (see section on Research Plans, below). Current work in the restricted domain of clinical trials will, we hope, illustrate general principles in the design of decision makers that gather expertise from written text and multiple knowledge sources.

D. Relevant Publications

Haggerty, J.: *REFEREE and RULECRITIC: Two prototypes for assessing the quality of a medical paper*. REPORT KSL-84-49. Master's Thesis, Stanford University, May 1984.

E. Funding Support

REFEREE currently receives only a small amount of funding. Most of the research is performed in time contributed by the researchers to this project.

Title: Knowledge-Based Systems Research

PI: Edward A. Feigenbaum

Agency: Defense Advanced Projects Research Agency

Grant identification number: N00039-86-0033

Total award period and amount: 10/1/85 - 9/30/88 $4,130,230 (in negotiation) (direct and indirect)

Current award period and amount: 10/1/86 - 9/30/87 $1,549,539 (direct and indirect) REFEREE component is $29,296, or 1.9 % of grant total.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

Dr. Brown and Dr. Feldman of the Stanford University School of Medicine are actively involved in the REFEREE project and are the primary domain experts for this project.

C. Critique of Resource Management

The SUMEX computer resource and Lisp workstations have been very important for the work to date, and the SUMEX staff has continued to be very cooperative with the REFEREE project.

III. RESEARCH PLANS

A. Goals & Plans

The overall objective of the REFEREE project is to use recent Artificial Intelligence techniques to build a system that helps the informed but statistically non-expert reader to evaluate critically the medical literature on randomized controlled trials (RCT's). This system will contain and be able to apply dynamically the detailed specialized knowledge of Dr. Byron W. Brown, a biostatistician expert in the design and evaluation of randomized controlled trials. We have divided our overall objective into two goals:

- Goal 1 is the construction of an expert system to help readers (e.g. medical students, medical researchers, clinicians, journal editors, or editorial assistants) assess the credibility of a *single* conclusion drawn from a *single* journal report of a randomized controlled trial. We have already made substantial progress toward this goal with the development of the prototype REFEREE system.

- Goal 2 is the expansion of REFEREE to an expert system that can be used by a similar range of readers to facilitate the evaluation of *multiple* reports based on randomized controlled trials. This expanded system, to be known as the REVIEWER, will thus perform meta-analysis.

The task of extending and refining the prototype REFEREE system in order to achieve these goals can be characterized in terms of three dimensions:

- Making the system more accessible to a variety of people by improving the user interface, validating the system's performance with different types of users, and providing an explanatory capability

- Expanding the knowledge base by continuing the knowledge acquisition process to cover additional types of RCT's

- Improving the inference engine to ensure consistency of the knowledge base and to focus the consultation process on questions relevant to the situation and the individual user.

The specific steps that are planned for the enhancement of the REFEREE system include the following:

1. Critique individual clinical trials according to the methodological quality of the trial;

2. Measure the efficacy of treatment as demonstrated in a randomized control trial;

3. Compare and contrast the credibility and efficacy of treatment reported by multiple journal articles; and

4. Combine the *qualitative* techniques of heuristic reasoning and the *quantitative* methods of statistical meta-analysis to extract a consensus opinion from multiple knowledge sources.

In addition, plans for Goal 2, the REVIEWER system to analyze multiple RCT's and form a consensus judgment, include:

1. Complete a review of the available literature on meta-analysis and augment the REFEREE prototype to produce estimators for meta-analysis and incorporate expert knowledge on the appropriateness of these methods.

2. Add explicit and heuristic knowledge needed for the calculation of robust, non-parametric estimators of effect size.

3. Construct a prototype of a system that builds categorical models in the domain of meta-analysis, to perform autonomous investigations in the domain of statistical model-building. The REVIEWER will utilize expert knowledge in biostatistics to guide its search for meaningful models.

4. Build a prototype of a system that can explore the domain of regression models for multiple RCT's that will use expert knowledge in its selection of predictor variables.

5. Package the REVIEWER in a form suitable for use by physicians and their assistants.

6. Verify the expertise of the REVIEWER system on a suite of papers drawn from clinical trials, similar to the validation of REFEREE above.

B. Justification for continued SUMEX use

The local area network maintained by the SUMEX staff is essential to the effective development and use of the REFEREE system on Lisp workstations. The availability of the Xerox workstations makes possible the evaluation of prototypes in that environment, and also facilitates the development of good user interfaces. The connections through the 2060 to local and national computer networks such as ARPAnet are important for sharing ideas and results with other medical researchers.

C. Need for other computing resources

The REFEREE project needs access to an additional high performance Lisp workstation to assist in the development and execution of the REFEREE programs. Such a machine is important to explore user interface issues, in addition to building the knowledge base for current and planned development. In addition, we intend to explore the implementation of REFEREE on less expensive personal computers such as the Macintosh II and other high performance machines. We anticipate the need for at least two of these machines for transporting our system and developing new modes of interaction with both naive and experienced users.

IV.D. 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.

IV.D.1. PATHFINDER Project

PATHFINDER Project

Bharat Nathwani, M.D.
Department of Pathology
University of Southern 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 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 University of Southern California and the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 72 common benign and malignant diseases of the lymph node based on 110 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.

PATHFINDER computer science research is focused on the exploration and extension of formal techniques for decision making under uncertainty. Research foci include (1) the assessment and representation of important probabilistic dependencies among morphologic features and diseases, (2) the representation of knowledge about the progression of disease over time, (3) the acquisition and use of independent expert knowledge bases, (4) the customization of the system's reasoning and explanation behaviors to reflect the expertise of the user, and, (5) the explanation of complex formal reasoning techniques.

Toward the pragmatic goal of constructing a useful pathology teaching and decision support system, PATHFINDER investigators are attempting to use intelligent computation to substantially increase the quantity and quality of pathology knowledge available to pathologists. Important areas of this knowledge integration task involve ongoing research on the crisp definition important morphologic features and feature severities, the synthesis of information from multiple experts, the translation among multiple pathology classification schemes, and the incorporation of knowledge about advances in immunology, cytogenetics, cell kinetics, and immunogenetics.

A group of expert pathologists from several centers in the U.S. have showed interest in the program and helped to provide the structure of the knowledge base for the PATHFINDER system.

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 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 ten 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 Previous Accomplishments

Since the project's inception in September, 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 by 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 original version of PATHFINDER is written in the computer language Maclisp and runs on the SUMEX DEC-20. This was transferred to Portable Standard Lisp (PSL) on the DEC-20, and later transferred to PSL on the HP 9836 workstations. Two graduate students, David Heckerman and Eric Horvitz, designed and implemented the program and are continuing to lead research on the project.

The prototype knowledge base was constructed by Dr. Nathwani. During the early part of 1984, we organized two meetings of the entire team, including the pathology 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.

During the last two years, we have focused on methodologies for more accurately representing expert beliefs. In particular, we have used *influence diagrams* to represent dependencies among features in the PATHFINDER knowledge base. A great deal of effort has been devoted to assessing and representing the intricate relationships among features that exist in the domain. We believe that this process will help to overcome some of the limitations of medical diagnostic systems.

We have also focused on the problem of complex information-theoretic inference. The explanation of a systems diagnostic behavior has been found to be of extreme importance to physicians. Unfortunately, it is often difficult to explain reasoning based on optimal models of inference. We have worked on the use of a set of alternative abstraction hierarchies to control inference. Our current techniques enable us to trade off optimality for the transparency of reasoning. We are now studying the control of this tradeoff to optimize inference.

C.1 The PATHFINDER knowledge base

The basic building block of the PATHFINDER knowledge base is the disease profile or *frame*. Each 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*.

Qualitative dependencies among features for each disease are represented using the influence diagram methodology mentioned above. An influence diagram contains *nodes* and *arcs*. Nodes represent features and arcs represent dependencies among features. In particular, an arc is drawn from one feature to another when an expert believes that knowing one feature can change his beliefs that another feature will take on its possible values even when the diagnosis is known. Probabilities are used to quantitate the beliefs asserted by the expert.

C.2 Hewlett-Packard Workstation

Through the USC-affiliated Information Sciences Institute, Dr. Nathwani has obtained a Hewlett-Packard Workstation that is similar to the 9836. The Pathfinder program has been brought up on this machine. This means that the program now exists on three different machines, in three separate locations, using one standard language (Portable Standard Lisp). Thus, the need for support of networked machines and communications has increased during this last year. Current plans are to move the system onto the Macintosh II system.

D. Publications Since January 1984

1. Horvitz, E.J., Heckerman, D.E., Nathwani, B.N. and Fagan, L.M.: *Diagnostic Strategies in the Hypothesis-directed PATHFINDER System, Node Pathology*. HPP Memo 84-13. Proceedings of the First Conference on Artificial Intelligence Applications, Denver, Colorado, Dec., 1984.

2. Heckerman, D. E., and Horvitz, E. J., "The Myth of Modularity in Rule-based Systems," in Uncertainty in Artificial Intelligence, Vol. 2, J. Lemmer, L. Kanal, ed., North Holland, New York, 1987.

3. Horvitz, E.J., Heckerman, D.E., Nathwani, B.N. and Fagan, L.M.: *The Use of a Heuristic Problem-solving Hierarchy to Facilitate the Explanation of Hypothesis-directed Reasoning*. KSL Memo 86-2. Proceedings of MedInfo, Washington D.C., October, 1986.

4. Horvitz, E. J., "Toward a Science of Expert Systems," Invited Paper, Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface, American Statistical Association, March, 1986, pgs. 45-52.

5. Heckerman, D.E., "An Axiomatic Framework for Belief Updates," in Uncertainty in Artificial Intelligence, Vol. 2, J. Lemmer, L. Kanal, ed., North Holland, New York, 1987.

E. Funding Support

Research Grant submitted to National Institutes of Health
Grant Title: "Computer-aided Diagnosis of Malignant Lymph Node Diseases"
Principal Investigator: Bharat Nathwani
Funding for three years from the National Library of Medicine
1 RO1 LM 04529
$766,053 (direct and indirect)

Professional Staff Association, Los Angeles County Hospital, $10,000.

University of Southern California, Comprehensive Cancer Center, $30,000.

Project Socrates, Univ. of Southern Calif., Gift from IBM of IBM PC/XT.

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 USC, 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. We have been in electronic contact and have met with members of the INTERNIST/CADUCEUS project, as well as been able to utilize information and experience with the INTERNIST program gathered over the years through the AIM conferences and on-line interaction. 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.

The SUMEX pilot project, RXDX, designed to assist in the diagnosis of psychiatric disorders, is currently using a version of the PATHFINDER program on the DEC-20 for the development of early prototypes of future systems.

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 knowledge base of the program is about to undergo revision by the experts, and then will be extensively tested. A logical next step would be to extend the program to clinical settings, as well as possible extensions of the knowledge base.

Other possible extensions include: developing techniques for simplifying the acquisition and verification of knowledge from experts, and creating 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 so that overlapping features are not incorrectly weighted in the decision making process, implementing new methods for scoring hypotheses, and creating appropriate explanation capabilities.

B. Requirements for Continued SUMEX Use

We are currently dependent on the SUMEX computer for the use of the program by remote users, and for project coordination. We have transferred the program over to Portable Standard Lisp which is used by several users on the SUMEX system. While the switch to workstations has lessened 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 Macintosh II workstations. We will need additional file space on the 2060 as we quadruple the size of our knowledge base through the construction of multiple knowledge bases. 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. Until the type of workstations that we are using in this research becomes inexpensive, we will continue to need a machine like SUMEX to provide others with a chance to experiment with our software.

IV.D.2. RXDX Project

RXDX Project

Robert Lindsay, Ph.D.
Michael Feinberg, M.D., Ph.D.
University of Michigan
Ann Arbor, Michigan

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 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 program uses a knowledge base constructed by encoding the clinical expertise of a skilled psychiatrist in a set of rules and other knowledge structures. 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 is a documented shortage of psychiatrists in the US (GMENAC, 1980), and the estimates of the prevalence of psychiatric illness used to develop that report were lower than the figures in recent population surveys (Myers et al., 1984). Further, most prescriptions for antidepressants are written by non-psychiatrists (Johnson, 1974; Kline, 1974) and the great majority of depressed patients seen by a sample of primary care physicians were treated inappropriately (Weissman et al., 1981). These data highlight the need for improving the treatment provided to the majority of mentally ill patients. We believe that computers can act as consultants to non-psychiatrist clinicians, resulting in improved patient care.

The potential benefits to psychiatry include: making relatively skilled psychiatric consultation widely available in underserved areas, including some public mental health facilities where patients are seen by non-psychiatrists and have relatively little direct patient-physician contact; providing non-psychiatrically trained physicians with additional information about psychiatric diagnosis and treatment; avoiding errors of oversight caused by inaccessible patient data; and increased productivity in patient care. Like any good consultant, the program will be able to teach the interested user, and can function as a teaching tool independent of direct clinical application.

C. Highlights of Research Progress

Our major project during the past year has been an expert system for the somatic treatment of (endogenous) depression, where somatic treatment includes antidepressant drugs, electroshock, and lithium. We are writing this system using KEE, an expert system shell generously donated by Intellicorp, running on a Xerox 1108 workstation. We have been able to incorporate the work we did earlier on SUMEX, either directly by transporting the rules or indirectly by using what we learned about building expert systems in general. The knowledge base includes information about the side effects of each of the drugs and about the physiological mechanisms of these side effects. This information allows us to predict drug interactions and the likelihood of occurrence of various side effects in a given patient, and to base explanations on knowledge of the underlying physiology. The knowledge base also includes specific information about drug regimens, about preventing and treating side effects, and about how to take all of this into account in selecting a drug and dosage regimen for the individual patient.

D. List of Relevant Publications

1. Feinberg, M. and Lindsay, R. K.: *Expert systems in Psychiatry and Psychopharmacology*. Psychopharmacol. Bull., 22, 1986, 311-316.

2. Lindsay, R. K.: Expert Systems in Psychiatric Diagnosis: Rule-Based Systems. Presented at MedInfo86, Washington, D.C.

3. Feinberg, M.: What Psychiatrists Can't Do. Presented at Medinfo86, Washington, D. C.

E. Funding Support

None.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaboration and Program Dissemination via SUMEX

We have established via SUMEX a community of researchers who are interested in AI applications in psychiatry. We also have used the message system to communicate with other AI scientists at SUMEX and elsewhere.

B. Sharing and Collaboration with other SUMEX-AIM Projects

During this past year we have had no occasion to engage in collaboration with other SUMEX-AIM Projects.

C. Critique of Resource Management

Our sole use of the system this year has been for communication. This has been very useful, but hampered by difficulties in matching the characteristics of various networks and terminals. This has made use of SUMEX, even for mail, awkward. It would be helpful to have some assistance with these problems.

III. RESEARCH PLAN

A. Project Goals and Plans

Our immediate objective is to develop expert systems that 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 that 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 M.D.'s as an adjunct to consultation with a human expert. We plan also to focus on problems of the user interface and the integration of this system with other databases.

B. Justification and Requirements for Continued SUMEX use

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. Although we have moved our system to local workstations, the communications capability of SUMEX will continue to be important.

We anticipate that our requirements for computing time and file space will continue at about the same low level for the next year.

C. Needs and Plans for Other Computing Resources

We anticipate that the need for additional computing power will continue to be met by local workstations.

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 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 feels quite limiting, now that the basic concepts of computer networking have become available and have proved their potential.

We expect that the role of the SUMEX-AIM resource will continue to evolve in the direction of increased importance of communication, including graphical information, electronic dissemination of preprints, and database and program access. The need for computer cycles on a large mainframe will diminish. We hope to have continued access to the system for communication, but do not anticipate continued use of it as a Lisp computation server.

If fees for using SUMEX resources were imposed, this would have a drastically limiting effect on the value of the system to us. Even if we had a budget to purchase such services, the inhibiting effect of having a meter running would cause us to make less use of it that we should. We have been conscious of the costs of the system and feel that we have not used it imprudently, even though we have not directly borne its costs.

IV.D.3. Dynamic Systems Project

Decision Support for Time-Varying Clinical Problems

Lawrence Widman, M.D., Ph.D.
Division of Cardiology
Case Western Reserve University
2065 Adelbert Road
Cleveland, OH 44106
(216) 844-3153

I. SUMMARY OF RESEARCH PROGRAM

A. PROJECT RATIONALE

Time-varying systems, which include many areas of medicine, science, economics, and business, can be described mathematically by differential equations. They are distinct from the pattern-matching and logic-based domains dealt with so successfully by existing expert system methods, because they can include feedback relationships. It is generally felt that they are best approached by enhancement of existing methods for deep model-based reasoning.

The goal of this project is to develop AI methods for capturing and using knowledge about time-varying systems. The strategy is to address general problems in model-based knowledge representation and reasoning. The intermediate objective is to develop methods which are powerful enough to work in selected realistic situations yet are general enough to be transportable to other, unrelated knowledge domains.

The tactical approach is to work on well-defined yet complex and interesting problems in the medical domain. We have, therefore, selected the human cardiovascular system as our prototype of a time-varying system, and are developing methods for representing and reasoning about its mechanical and electrical activities in the normal and diseased states.

A.1 Technical Goals

This project presently has two distinct tracks: hemodynamic modeling and cardiac arrhythmia interpretation.

1. Hemodynamic Modeling

The goals of this subproject are to develop:

(a) a knowledge-representation method using symbolic modeling which captures the qualitative and, when possible, the quantitative behavior of systems with feedback relationships. Preferably, the symbolic model should be translatable into the differential equations which describe the behavior of the system being modeled.

(b) a reasoning method based on the symbolic modeling tool created in subgoal (a) which permits the inference of differential diagnoses (a set of hypothesized diagnoses) from incomplete data.

(c) a reasoning method based on subgoals (a) and (b) which permits inference of the state of the model for each hypothesized diagnosis. This subgoal would be satisfied by an algorithm which specifies a self-consistent set of values for all variables in the model, for a given hypothesis based on a given set of data. Such sets of data would constitute initial conditions for differential equations derived from the model.

(d) a simulation method, based on the model and its equivalent differential equations together with the initial conditions derived from the differential diagnosis (steps a-c above), for predicting the expected time course of the system being modeled for each hypothesized diagnosis. This method could also be used to predict the effects of treatments being considered for recommendation by the program.

(e) a reasoning method, based on domain-independent properties of the model, for shrinking and/or expanding the model automatically to use a minimal model configuration to account for normal and abnormal data.

(f) an explanation facility for examining the model, the given data, the inferred hypothesized diagnoses, predicted behaviors, and modifications of the model, to answer user queries and to teach fundamental concepts.

2. Cardiac Arrhythmia Recognition

The goals of this subproject are to develop:

(a) a symbolic model of the electrical system of the human heart, including pertinent anatomic and electrophysiologic features of the normal and diseased heart. The electrophysiologic features would include deterministic characteristics (e.g., conduction velocities, refractory periods), stochastic features (e.g., behavior of automatic foci), and temporal interactions (e.g., competing pacemakers).

(b) a symbolic/numeric representation of the observable features of the electrical activity of the heart, both surface EKG and intracardiac recordings, including noise. This representation would be intended to allow a feature extraction module working on actual patient data to communicate with a symbolic reasoning module, and would be translatable directly into waveform display format.

(c) a reasoning method for extracting features from raw, digitized signal data. This method would augment established signal processing techniques by using knowledge-based algorithms to improve detection of P and T-U waves and to improve rejection of noise. It should be noted that this is itself a major research undertaking in the signal processing domain.

(d) a reasoning method for inferring the cardiac rhythms consistent with a given disease state in the model, similar to the prediction of consequences of the hemodynamic model in the first subproject. The output of this method would be in the symbolic/numerical representation of subgoal (b).

(e) a reasoning method for inferring possible disease states in the model from a given feature-extracted recording of the electrical activity of the heart. This subgoal constitutes cardiac arrhythmia interpretation, and is itself a major research project.

(f) a categorization method for inferring hierarchies of diagnoses from elementary abnormalities. For example, "periods of atrial fibrillation up to 30 minutes at up to 150 beats/min, supraventricular tachycardia of up to 10 beats length at a rate of 130 beats/min, and sinus bradycardia with a minimum rate of 45, all consistent with the sick sinus ("tachy-brady") syndrome" and "two QRS morphologies are present: they are narrow at rates less than 120 and are wide at rates above 120, consistent with a rate-dependent bundle branch block".

(g) an explanation facility for examining the model, the input data, and the interpretations to answer user queries and to teach fundamental concepts.

B. MEDICAL RELEVANCE AND COLLABORATIONS

The two subprojects have related but separate medical goals:

1. Hemodynamic Modeling.

There are three subgoals in this subproject: model-based sensor integration, model-based caregiver assistance, and model-based experiment interpretation.

a. Model-based Sensor Integration.

The long-range application of this subproject is the integration of patient-related data in the intensive care environment. Model- based real-time systems would allow the system to share a global understanding of the patient's condition with the human caregivers. Thus, it could interpret significant trends in key parameters and could draw attention to relationships which might otherwise escape attention in the constant flood of data common to these environments.

b. Model-based Caregiver Assistance.

It could also serve as an assistant to the caregiver. In this mode, the human caregiver could evaluate the merits of proposed diagnostic and therapeutic measures in light of available data on the patient's condition.

Practical application of these concepts requires further development of the model and the reasoning algorithms, and extensive testing against real clinical scenarios. Refinement and quality control are presently the responsibility of the principal investigator, who is a board-certified internist with subspecialty training in invasive cardiology.

Practical application also awaits gen . acceptance of standardized hospital data buses for automatic acquisition of importa... parameters now stored primarily on paper or on computers outside the intensive care setting, such as fluid inputs and outputs, medications, and results of invasive and non-invasive tests. Further, improved user interfaces will require better graphics and increased computer literacy on the part of caregivers.

c. Model-based Experiment Interpretation.

An intriguing third application of this subproject is in the area of interpretation of biomedical experiments. The symbolic model concept, which enforces objectivity, can assist investigators by allowing them to compare alternate interpretations of experimental data. In this application, several alternate models would be proposed by the experimenter to explain a given experimental outcome. The consequences of each model given different experimental parameters could then be evaluated and compared with real data to confirm or refute competing proposed models.

The advantage of using a computer in this manner is the guarantee of self-consistent and objective exploration of each possibility. The advantage of using a symbolic model, rather than a numerical model such as the Guyton-Coleman model or a simpler derivative, is that the underlying cause and effect relationships are explicit and can be easily modified by the experimenter. The AI interest in this subgoal would be the refinement of the symbolic model through application to real experiments. Unlike in the MOLGEN project at Stanford, automatic hypothesis formation would not be an objective in this subgoal.

A new collaboration to explore this application is being explored with Dr. E. Merrill Adams, an experimental physiologist in the Department of Surgery at Case Western Reserve University School of Medicine. Dr. Adams approached us because of his long-standing interest in applying AI techniques to his experiments on the interactions of the cardiovascular and pulmonary systems. A discussion group is being organized, and we hope to continue despite the move of the principal investigator to Texas this summer.

2. Cardiac Arrhythmia Recognition.

The long-range application of this sub-project is in clinical devices such as intensive-care arrhythmia monitors, portable Holter monitors, and implantable cardioverter-defibrillators. There are two subgoals: recognition of surface electrocardiographic (EKG) recordings and recognition of intracardiac recordings.

a. Recognition of surface electrocardiographic recordings.

Substantial and well-recognized obstacles in signal processing will likely prevent non-AI algorithms from advancing beyond the current state of the art of interpretation of surface EKG recordings. These obstacles are primarily the problems of reliable detection of P and T-U waves, and rejection of noise. We hope that AI techniques will be helpful with these problems, as is suggested by the work of Muldrow et al. (Computers and Cardiology, 1986, in press).

We hope further that, by mimicking the behavior of expert human cardiologists, these obstacles can be bypassed if they cannot be overcome. We have enlisted as consultants Dr. William Long of M.I.T., who supervised Muldrow in the paper cited above, and Dr. Benjamin Kuipers of the University of Texas at Austin, who is interested in AI techniques for physiological modeling.

b. Recognition of intracardiac recordings.

Intracardiac recordings, which are taken from wires placed in the heart by percutaneous venous puncture or around the heart by surgery, are relatively free of P wave ambiguity and of noise. They are representative of the quality of signals available to implantable cardioverter-defibrillators.

Cardioverter-defibrillators are devices like pacemakers in that they monitor the heart rhythm in a patient to determine if an abnormality exists. They are capable of taking action (electrical countershock) if an appropriate abnormality is detected. Unlike ordinary pacemakers, these devices detect abnormalities characterized by rapid rates of heart activity, rather than excessively slow rates. They have been shown to reduce one-year mortality in high-risk patients from 30% to 2%, and they are expected to play an increasingly large role in treatment of such patients.

These relatively new devices currently use quite simple algorithms to detect abnormalities. The action they take consists of applying an electrical shock directly to the heart. This shock is frequently unpleasant to the patient. The problem is that the algorithms sometimes confuse innocent rapid heart rates, such as from exercise or atrial tachyarrhythmias, with lethal ventricular arrhythmias. This has proved troublesome enough to prompt repeated calls in the electrophysiology literature for improved algorithms for arrhythmia recognition in these devices.

The algorithms developed in this subproject would be suitable for this application when the computer power in the devices improves. Because these devices require powerful energy sources to perform repeated shocks over their lifetimes of 2-3 years, the power drain of more sophisticated computer chips is less important than it would be in ordinary pacemakers.

C. Highlights of Research Progress

1. Hemodynamic Modeling

Subgoals (a) through (d) have been accomplished in prototype form. The approach relies on a semi-quantitative representation [subgoal (a)] which assigns values by default if the user does not specify more detailed information. The second phase of this project yielded subgoal (d), the simulation of a given model. This phase was accomplished by translating the model into a set of dynamical systems equations, which were then integrated in the standard manner.

More recently, subgoals (b) and (c) have been accomplished in prototype form. Constraint propagation using a dynamically generated semi-quantitative quantity space is performed by interpreting the model as a set of constraint equations. Domain-independent heuristics which recognize morphological features of the model are used to further constrain the propagation of constraints and to generate hypotheses when ambiguities arise. These heuristics generate a set of self-consistent hypotheses, each of which is a hypothesized diagnosis (subgoal b). Dr. Yong-Bok Lee of the Case Western Reserve University Department of Electrical Engineering and Applied Physics participated in this subgoal for his doctoral dissertation. The doctoral dissertation, awarded in August, 1986, was co-supervised by Professor Yoh-Han Pao of that Department and by Dr. Widman.

Each hypothesized diagnosis is then refined by mathematical relaxation, in which the propagated values are treated as initial guesses, and the values are refined iteratively, again by interpreting the model as a set of constraint equations (subgoal c). In the several scenarios which have been examined, the value assignments achieved by hypothesis and iterative refinement have achieved correlation coefficients up to 0.90 with the values obtained by simulation of the same model.

We do not anticipate beginning work on the remaining subgoals until the above prototype methods have been further refined and tested.

2. Cardiac Arrhythmia Recognition

This subproject is just beginning. We have built a prototype symbolic model of the electrical conduction system of the heart and have reproduced simple rhythms. The important issues of stochastic variation and of noise have not been addressed. We are hopeful that important insights will obtained from newly developing literature on stochastic simulation (e.g., Pearl, J.: *Evidential Reasoning Using Stochastic Simulation of Causal Models*. Artificial Intelligence. 1987;32:245-257).

Following the move of the principal investigator to Texas, this subproject will replace the hemodynamic modeling subproject as the major research focus. This research effort will be supported in part by a Grant-in-Aid from the American Heart Association, Texas Affiliate.

The principal investigator will have access to intracardiac signals from a variety of appropriate patients on the clinical service at his hospital complex. This should facilitate the development of practical algorithms.

We have also begun discussions with a major pacemaker manufacturer with the goal of establishing a working relationship. The purpose of the relationship would be to enable practical pacemaker manufacturing constraints to be taken into account from an early stage in the development of this subproject. So far, the discussions have demonstrated interest on both sides, but will require further algorithm development in order to proceed.

D. List of Relevant Publications

1. Widman, L.E. Reasoning about Diagnosis and Treatment in a Causal Medical Model using Semi-Quantitative Simulation and Inference. Workshop on Artificial Intelligence in Medicine, National Conference on Artificial Intelligence AAAI-87, Seattle.

2. Widman, L.E., Lee, Y.-B., and Y.-H. Pao. Diagnosis of Causal Models by Semi-Quantitative Reasoning. (submitted to SCAMC 1987).

3. Widman, L.E., Lee, Y.-B., and Y.-H. Pao. Diagnosis of Causal Medical Models by Semi-Quantitative Reasoning. In: Miller, P.L. (ed.). Topics in Medical Artificial Intelligence, Springer-Verlag (in preparation).

4. Lee, Y.-B. and L.E. Widman. Reasoning about Diagnosis and Treatment in a Causal Time-varying Domain using Semi-Quantitative Simulation and Inference. Workshop on Artificial Intelligence and Simulation, National Conference on Artificial Intelligence AAAI-86, Philadelphia.

5. Widman, L.E. Representation Method for Dynamic Causal Knowledge Using Semi-Quantitative Simulation. Fifth World Conference on Medical Informatics. 1986: 180-184.

E. Funding Support

1. American Heart Association, Texas Affiliate
Grant-in-Aid Award.
Knowledge-Based Computer Algorithms for Arrhythmia Analysis.
Principal Investigator: Lawrence E. Widman.
Award period: July, 1987 - June, 1988.
Level: $24,850 direct costs.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Sharing and interactions with other SUMEX-AIM projects

The major interactions with SUMEX-AIM have been (1) computational support and (2) communication with members of the AIM community.

(1) SUMEX-AIM is the major source of computing power at this time. Dr. Widman expects that a LISP workstation will be available after he relocates to Texas this summer. SUMEX-AIM computing power will then be needed primarily for demonstrations at meetings and as backup during workstation down-time.

(2) SUMEX-AIM is the current electronic mailbox. Its central location allows ready Email access by users of Arpanet, Bitnet and Csnet. This access has proved invaluable to Dr. Widman in communicating rapidly and effectively with co-workers at other institutions. The value of this type of communication has been demonstrated several times during the past year, when he had to make major career, equipment negotiation, and manuscript revision decisions, without local expertise, within short periods of time.

Review of the longer term history of this project shows that it would not exist had SUMEX-AIM not provided telecommunication support for the initial feasibility project in 1984-1985, which was carried out on the computers of the MIT Laboratory for Computer Science, Clinical Decision Making Group.

C. Critique of Resource Management

The service provided by SUMEX-AIM has been exemplary, largely because of prompt and effective response to difficulties as they arise. There has been a clear effort to assure that telecommunication access remained reliable during changes in commercial vendors, and the staff have responded to several technical questions promptly and accurately. Down-time has been minimal compared to that of other systems we have used, and is almost always scheduled several days in advance.

The reason we sought contact with the AIM community was that it seemed the natural niche for our research interests. There is no short-term prospect that this project will reach commercial maturity or that it will lose sight of fundamental AI issues, and so we feel that it still belongs in the scientific AIM framework.

As noted in the previous section, the communication with other members of the AIM community has proved invaluable in the advancement of this project.

III. RESEARCH PLANS

Project Goals and Plans

The long range goals of this project are to develop intelligent comprehensive monitoring/alarm systems for intensive care unit settings; and intelligent arrhythmia recognition systems for monitors, Holter recorders, and implantable cardioverter/ defibrillators. The short term strategies for achieving these goals are discussed above.

The next phase of this research will be conducted at the University of Texas Health Science Center at San Antonio. Dr. Widman will be joining the faculty of Medicine there on July 1, 1987, in the Division of Cardiology. His clinical duties will include invasive hemodynamic and electrophysiological studies on selected patients. Substantial time is committed to research, and this project will constitute his major research emphasis.

B. Justification and requirements for continued SUMEX use

The justification for this project is its potential for advancing the state of the art of expert system technology in the area of temporal reasoning and deep causal modeling, and for demonstrating practical use of expert symbolic computing in potentially life-saving, knowledge-intensive environments.

The requirements for continued SUMEX-AIM use should be the same as currently: telecommunications support, Arpanet access, about 3 megabytes of disc space, and a reasonable amount of CPU time. When the Lisp workstation becomes available (see below), the requirement for telecommunication support and CPU time should decrease.

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

The symbolic computing needs for the hemodynamic modeling subproject are being met by SUMEX. Once embarked on the arrhythmia recognition subproject, there will be a strong need for high-resolution graphics, and processing of tens of megabytes of data. To meet these needs, a Lisp workstation will be provided by the University of Texas. Data acquisition in real time and initial signal processing will be done with an IBM AT class microcomputer equipped with a standard third-party multichannel analog-to-digital converter. Communication between the machines will be by RS232 or the local Ethernet LAN. Once these machines are in place, SUMEX-AIM will be needed primarily for communication and demonstration projects, as noted above.

D. Recommendations for Future Community and Resource Development

Our strong recommendation is that SUMEX-AIM be maintained as a national AIM resource for communication, development of software useful to the AIM community, and sharing of demonstration projects. SUMEX-AIM could also serve as a central source of advice for new workstation users who may be geographically isolated from experienced workstation users.

Additionally, we would strongly support retention of the current telecommunication support and enough computing power to support promising young investigators who would otherwise not have access to symbolic computing power.

IV.D.4. Knowledge Engineering for Radiation Therapy

KNOWLEDGE ENGINEERING FOR RADIATION THERAPY

Ira J. Kalet, Ph.D.
Witold Paluszinski
University of Washington
Seattle, Washington

I. Summary of Research Program

A. Project Rationale

We are developing an expert system for planning of radiation therapy for head and neck cancers. The project will ultimately combine knowledge-based planning with numerical simulation of the radiation treatments. The numerical simulation is needed in order to determine if the proposed treatment will conform to the goals of the plan (required tumor dose, limiting dose to critical organs). The space of possible radiation treatments is numerically very large, making traditional search techniques impractical. Yet, with modern radiation therapy equipment, the design of treatment plans might be significantly aided by automatically generating plans that meet the treatment constraints. The project will result in systematization of knowledge about radiation treatment design, and will also provide an example of how to represent and solve design problems with a knowledge based system.

B. Medical Relevance and Collaborations

Radiation therapy has shown dramatic improvement in the cure rate for many tumor sites in the last two decades. Much of this can be attributed to the improved penetration capability of modern megavoltage X-ray machines. These high energy beams can deliver high tumor doses without overdosing surrounding tissue in many cases. However, they are typically used in very limited ways, because of the lack of suitable simulation systems to compute the dose distribution for any but a few narrow choices of treatment geometry. In the last few years these simulation systems have been extended to the full range of geometric treatment arrangement that any therapy machine is capable of. Thus it would be valuable to be able to generalize our knowledge of treatment technique by exploring these expanded possibilities. In addition, even treatments with standard geometries can be very complex, and it is tedious to explore all of them individually. A knowledge-based system can generate a few "best" plans which satisfy the constraints and allow more time for the physician to evaluate the options, or make minor adjustments for optimization.

Since cancer treatment is a multi-disciplinary approach involving surgery and chemotherapy as well as radiation, it is important to coordinate this work with knowledge-based program projects in those areas. Most significant is the ONCOCIN project, which addresses management of patients on chemotherapy protocols.

This project has some relevance to computer science as well, in that our approach, if successful, may contribute to a better understanding of design problem solving with knowledge-based systems.

C. Highlights of Research Progress

In the past year, we have made significant additions to the rule database for details of head and neck cancer treatment. We have devised a representation of parameters for radiation treatment fields and created a set of prototype treatment field arrangements. The prototypes are used as building blocks for constructing complex treatment plans. In addition we have examined the issues of control strategy associated with using prototypes in planning.

Our expert system now has about two hundred rules, a two-level (agenda-based) control strategy, and about ten prototypes for plan construction. It is written in Interlisp on a VAX running the VMS operating system. This environment was chosen because it is also the environment used for a graphic simulation system that does radiation dose calculations for arbitrary treatment plans. The dose calculation is needed to determine whether a plan meets the treatment goals set by the system in its early phases of planning.

D. List of Relevant Publications

1. I. Kalet and W. Paluszynski: A Production Expert System for Radiation Therapy Planning. Proceedings of the AAMSI Congress 1985, May 20-22, 1985, San Francisco, California. Edited by Allan H. Levy and Ben T. Williams. American Association for Medical Systems and Informatics, Washington, D.C., 1985.

2. W. Paluszynski and I. Kalet: Radiation Therapy Planning: A Design Oriented Expert System. WESTEX-87 (Western Conference on Expert Systems), Anaheim, California, June 2-4, 1987.

3. I. Kalet and J. Jacky: Knowledge-based Computer Simulation for Radiation Therapy Planning. Proceedings of the Ninth International Conference on the use of Computers in Radiotherapy, Scheveningen, the Netherlands, June 1987. North Holland, 1987.

II. Interactions with the SUMEX-AIM Resource

Our main use of the SUMEX-AIM resource has been as a means to be in contact with other researchers working on AIM projects. The existence of a mailbox at SUMEX-AIM has made it much easier for colleagues at other institutions to communicate with us, and has been valuable in assisting us with organizing the AIM Workshop for 1987.

We have had a great deal of contact with members of the ONCOCIN project and other groups. This has been valuable to us in stimulating creative approaches to our project.

III. Research Plan

A. Project Goals and Plans

We plan to continue to acquire rules and develop our current expert system. This includes solving problems of use of prototypes, satisfaction of constraints by some kind of backtracking search, and incorporating evaluation of plans by using the results of the dose computation. This last idea involves coupling the expert system with the dose computation system (written in PASCAL) in suitably efficient ways. Our long-term goal is to shape the user interface and improve the system performance to where it can provide assistance to clinicians in treatment design for patients in the normal course of treatment.

B. Justification and Requirements for Continued SUMEX use

We foresee continued need to be in touch with other members of the AIM community, particularly projects centered at SUMEX. While we do not expect to use the computing resources of SUMEX directly, some more extensive communication and involvement is likely to be useful.

C. Plans For Other Computing Resources

The main computing resources for our project will continue to be local. We will be rewriting the expert system code in VAX Lisp, an implementation of Common Lisp on the DEC VAXstation. We expect delivery of a VAXstation II/GPX in the near future. This appears to be a good choice to satisfy our need for high performance graphic simulation and a reasonable Lisp system. However, the resources for the dose computation may not be adequate as we incorporate more sophisticated computation models. As this develops, we hope to experiment with distributed systems, in which the dose computation may run on a remote resource, which may or may not be at SUMEX.

D. Recommendations for Future Community and Resource Development

Two areas will be of increasing importance to us in the future: communication capabilities (electronic mail and file transfer) and centralized databases. By centralized databases, we refer to the need for better maintenance of mailing lists, information about projects, and possibly on-line reports. Dr. Kalet's experience in organizing the AIM Workshop for 1987 demonstrated that electronic communication is invaluable, even in its present state, but in order to create a list to send announcements to, we expended many hours of manually cutting and pasting messages containing past lists and searching for up-to-date electronic mail addresses.

If fees for use of SUMEX resources were imposed, the main impact on our project would be one of increased isolation, unless we could find grant support for the fees.

IV.D.5. Pathophysiologic Diagnosis Project

COMPUTER-BASED EXERCISES IN PATHOPHYSIOLOGIC DIAGNOSIS

J. Robert Beck, M.D.
Dartmouth College School of Medicine
2 Maynard St.
Hanover, N.H. 07355

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

Research in artificial intelligence at Dartmouth Medical School focuses on three main areas: 1) knowledge-based systems applied to laboratory medicine and pathology, 2) knowledge acquisition using machine learning techniques, and 3) computer-based instruction using artificial intelligence techniques to critique students' workup plans. These projects have in common the fundamental research questions of how knowledge should be represented and used in a classification approach to problem-solving related to the use of laboratory data.

Knowledge-based systems in laboratory medicine:

We are investigating the use of knowledge-based systems to review requests for blood products. A system is being developed to advise pathologists and pathology residents about the appropriateness of transfusion requests.

The system will have both diagnostic and therapeutic objectives. The diagnostic part of the system will be used to evaluate information available in machine-readable form, and then ask the user a few relevant questions. Based on the available information, the system will determine possible diagnoses relevant to transfusion medicine. The current prototype is focusing on coagulopathies and bleeding disorders. The objective is to have a system that can quickly provide a summary of relevant laboratory information to the pathologist charged with the responsibility of evaluating appropriateness of transfusion requests. The therapeutic recommendations of the system will be focused on determining appropriate choices and quantities of blood products or substitutes. One of the purposes of this investigation is to determine whether a knowledge-based system can eventually reduce inappropriate use of blood products. The purpose of the tool is not to usurp decision-making, but to pre-process large volumes of transfusion requests and large volumes of data on each request, in order to focus the pathologist's attention in a time-efficient manner on the most relevant information. The system is in the early knowledge acquisition stage. The initial prototype is being built using IBM's Expert System Environment tool.

Knowledge acquisition for knowledge-based systems:

The purpose of this project is to develop machine learning tools that can be used for knowledge acquisition from databases. The focus is on deriving classification rules in the form of criteria tables. The criteria table format has been used for many years in medicine, and is still in use particularly in the area of rheumatic diseases (for example, the ARA criteria for systemic lupus erythematosus). Other diseases for which diagnostic criteria tables have been developed include polycythemia vera, multiple myeloma and primary biliary cirrhosis. In addition, criteria tables have been found useful as a knowledge representation for expert systems.

We have developed a program, called the CRiteria Learning System (CRLS) which is capable of automatically generating criteria tables from a database of positive and negative examples. CRLS is implemented in Common LISP on a SUN-3 workstation. It utilizes not only the raw data but also some background knowledge supplied by the user about the concepts to be learned, the features of the problem, and the type of diagnostic performance the user wishes to optimize (i.e. sensitivity, specificity, efficiency, etc.). CRLS learns decision rules that are more comprehensible than the rules generated by other machine learning programs. Tests of the system have also shown that it is capable of handling large databases containing as many as 1500 cases with 50 variables each.

Teaching medical pathophysiology using computer-based tools:

The project "Computer-based Exercises in Pathophysiologic Diagnosis" is funded through the National Library of Medicine's Medical Informatics research initiative. It has four specific aims:

1. To develop two computer-assisted laboratory exercises for basic content areas (anemia and coronary artery disease) in second-year medical education, oriented toward the processes of diagnosis and evaluation, utilizing techniques of medical decision science, critiquing, and software engineering (the PLAN-ALYZER system);

2. To utilize the computerized teaching modules to test the hypothesis that students with access to process-oriented educational tools can integrate their didactic knowledge more effectively than with access only to non-process oriented traditional education, including lecture notes, texts, and non-intelligent audiovisual aids;

3. To develop practical application versions of the two PLAN-ALYZER models that can be used as diagnostic tools with more senior medical students, residents, and physicians in the clinic, providing them with a decision analysis tool and expert critiques of their evaluations of real patients;

4. To utilize the originally proposed and the advanced systems to explore the process of how the effective physician solves clinical problems, a process which has been found to be different from traditional problem solving.

PLAN-ALYZER prototyping is being accomplished on the Macintosh Plus and Macintosh II workstations, using the Macintosh Programmer's Workshop. Novel AI features of the PLAN-ALYZERs include a scoring metric based on unate boolean functions, to compare students' decision trees with gold standard trees, a mechanism by which augmented transition network critiques can be developed for decision models, and the encoding of the domain experts' instructional styles as well as content into the models.

An interdisciplinary team of computer scientists, physicians, and educators is working on the Computer-based Exercises project. A prototype system is nearing completion, with formative evaluation scheduled for Fall, 1987.

D. Relevant Publications

Beck, J.R., Prietula, M.J., Russo, E.A.: A role for intelligent systems in teaching medical pathophysiology. In: Salamon, R., Blum, B., Jorgenson, M. (eds). Proc. Fifth Conf. Med. Inform. (MEDINFO '86), Elsevier-North Holland, Amsterdam, 1986, 936-938.

Beck, J.R.: Artificial intelligence: A topic for Medical Decision Making? (edit.) Med. Decis. Making 1987; 7:4.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

The Dartmouth group is pleased to be a new addition to the SUMEX research resource. Most of our projects take place on the Dartmouth campus, but we require access to the national AI community in order to share ideas, disseminate research results, and grant our trainees and junior faculty access to the developments of others. Also, inasmuch as our research in medical educational applications of computer science and decision making has significant potential for dissemination, the SUMEX community of scholars forms a natural group for focusing and broadening our research ideas.

Appendix A
AIM Management Committee Membership

Following are the current membership lists of the various SUMEX-AIM management committees:

AIM Executive Committee:

SHORTLIFFE, Edward H., M.D., Ph.D. (Chairman)
Principal Investigator - SUMEX
Medical School Office Building, Rm. X271
Stanford University Medical Center
Stanford, California 94305
(415) 723-6970

FEIGENBAUM, Edward A., Ph.D.
Co-Principal Investigator - SUMEX
Heuristic Programming Project
Department of Computer Science
701 Welch Road, Building C
Stanford University
Stanford, California 94305
(415) 723-4879

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

LINDBERG, Donald A.B., M.D. (Past Adv Grp Chrmn)
Director, National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 20814
(301)496-6221

MYERS, Jack D., M.D.
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 648-9933

AIM Advisory Group:

MYERS, Jack D., M.D. (Chairman)
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 648-9933

AMAREL, Saul, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-3546

COULTER, Charles L., Ph.D. (Exec. Secretary)
Bldg 31, Room 5B41
Biomedical Research Technology Program
National Institutes of Health
9000 Rockville Pike
Bethesda, Maryland 20892
(301) 496-5411

FEIGENBAUM, Edward A., Ph.D. (Ex-officio)
Co-Principal Investigator - SUMEX
Heuristic Programming Project
Department of Computer Science
701 Welch Road, Building C
Stanford University
Palo Alto, California 94305
(415) 723-4879

KULIKOWSKI, Casimir, Ph.D.
Department of Computer Science
Hill Center Busch Campus
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

LINDBERG, Donald A.B., M.D.
Director, National Library of Medicine
Building 38, Rm. 2E-17B
8600 Rockville Pike
Bethesda, Maryland 20814
(301) 496-6221

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 20892
(301) 496-1168

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)
Principal Investigator - SUMEX
Medical School Office Building, Rm. X271
Stanford University Medical Center
Stanford, California 94305
(415) 723-6979

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) 723-4879

LEVINTHAL, Elliott C., Ph.D.
Departments of Mechanical and Electrical Engineering
Building 530
Stanford University
Stanford, California 94305
(415) 723-9037

SHORTLIFFE, Edward H., M.D., Ph.D.
Principal Investigator - SUMEX
Medical School Office Building, Rm. X271
Stanford University Medical Center
Stanford, California 94305
(415) 723-6979

Appendix B
Scientific Subproject Abstracts

The following are brief abstracts of our collaborative research projects.

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) 723-1997 (CLANCEY@SUMEX-AIM)

Bruce G. Buchanan, Ph.D.
Computer Science Department
701 Welch Road
Stanford University
Palo Alto, California 94304
(415) 723-0935 (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 consultation 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, being developed on D-machines.

REFERENCES

Clancey, W.J.: **Knowledge-Based Tutoring: The GUIDON Program**, Cambridge: The MIT Press, 1987.

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. 1984. (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) 725-3815

Contact: Dr. Peter FRIEDLAND@SUMEX-AIM
(415) 723-3728

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. Our current research concentrates on scientific discovery within the subdomain of regulatory genetics. We desire to 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.

SOFTWARE AVAILABLE ON SUMEX

SPEX system for experiment design.
UNITS system for knowledge representation and acquisition.
SEQ system for nucleotide sequence analysis.

REFERENCES

1. Friedland, P.E.: *Knowledge-based experiment design in molecular genetics*, (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-79-771.

2. Friedland, P.E. and Iwasaki, Y.: *The concept and implementation of skeletal plans,* Journal of Automated Reasoning, 1(2):161-208, 1985.

3. Friedland, P.E. and Kedes, L.: *Discovering the secrets of DNA,* Communications of the ACM, 28(11):1164-1186, November, 1985.

4. Stefik, M.J.: *An examination of a frame-structured representation system,* Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 845-852.

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

6. Karp, P., and D. Wilkins: *An Analysis of the Deep/Shallow Distinction for Expert Systems.* Stanford University Knowledge Systems Laboratory Report KSL-86-32, 1986.

7. Karp, P., and P. Friedland: *Coordinating the Use of Qualitative and Quantitative Knowledge in Declarative Device Modeling.* Stanford University Knowledge Systems Laboratory Report KSL-87-09, 1987.

8. Round, A.: *QSOPS: A Workbench Environment for the Qualitative Simulation of Physical Processes.* Stanford University Knowledge Systems Laboratory Report KSL-87-37, 1987.

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
Medical School Office Building
Stanford, California 94305
(415) 723-6979 (SHORTLIFFE@SUMEX-AIM)

Project Director: Dr. Lawrence M. Fagan (FAGAN@SUMEX-AIM)

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 limited experimental 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 generally been well-accepted since its introduction, and we are now testing a version of the program which runs on 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 clinical use but requires Lisp machines to be run. 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

1. 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.

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*. Int. J. Man-Machine Studies 19:479-496, 1983.

4. 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.

5. Hickam, D.H., Shortliffe, E.H., Bischoff, M.B., Scott, A.C., Jacobs, C.D.: A study of the treatment advice of a computer-based cancer chemotherapy protocol advisor (Memo KSL-85-21). *Annals of Internal Medicine* 103(6 pt 1):928-936 (1985).

6. Kent, D.L., Shortliffe, E.H., Carlson, R.W., Bischoff, M.B., Jacobs, C.D.: Improvements in data collection through physician use of a computer-based chemotherapy treatment consultant (Memo KSL-85-22). *Journal of Clinical Oncology* 3:1409-1417, 1985.

7. Tsuji, S. and Shortliffe, E.H.: Graphical access to a medical expert system: I. Design of a knowledge engineer's interface (Memo KSL-85-11). Meth. Inf. Med., 25:62-70, 1986.

8. Lane, C.D., Differding, J.C., Shortliffe, E.H.: Graphical access to a medical expert system: II. Design of an interface for physicians (Memo KSL-85-15). Meth. Inf. Med., 25:143-150, 1986.

9. Shortliffe, E.H. Medical expert systems: Knowledge tools for physicians. Memo KSL-86-52. Special issue on Medical Informatics, West. J. Med. 145:830-839, 1986.

Stanford Project: PROTEAN Project

Principal Investigators: Oleg Jardetzky
(JARDETZKY@SUMEX-AIM.STANFORD.EDU)
Nuclear Magnetic Resonance Lab, School of Medicine
Stanford University Medical Center
Stanford, California 94305

Bruce G. Buchanan, Ph.D.
(BUCHANAN@SUMEX-AIM.STANFORD.EDU)
Computer Science Department
Stanford University
Stanford, California 94305

Contact Person: Bruce G. Buchanan

The goals of this project are related both to biochemistry and artificial intelligence: (a) use existing AI methods to aid in the determination of the 3-dimensional structure of proteins in solution (not from x-ray crystallography proteins), and (b) use protein structure determination as a test problem for experiments with the AI problem-solving structure known as the Blackboard Model. Empirical data from nuclear magnetic resonance (NMR) and other sources may provide enough constraints on structural descriptions to allow protein chemists to bypass the laborious methods of crystallizing a protein and using X-ray crystallography to determine its structure. This problem exhibits considerable complexity, yet there is reason to believe that AI programs can be written that reason much as experts do to resolve these difficulties. A prototype knowledge-based system assembles major secondary structures of a protein into families of structures compatible with a given set of distance constraints under the control of an explicit assembly strategy. Structures can also be refined at the atomic level of detail using constraints within secondary structures and between amino acid side chains to further restrict the 3-dimensional structure found. By generalizing this approach to the assembly of arrangements of objects subject to constraints, we have developed a language for specifying actions and control for problem solving in similar problem domains.

REFERENCES

1. Altman, R. and Jardetzky, O.: *New strategies for the determination of macromolecular structures in solution*. Journal of Biochemistry (Tokyo), Vol. 100, No. 6, p. 1403-1423, 1986.

2. Altman, R. and Buchanan, B.G.: *Partial Compilation of Control Knowledge*. To appear in: Proceedings of the AAAI 1987.

3. Brinkley, J., Cornelius, C., Altman, R., Hayes-Roth, B. Lichtarge, O., Duncan, B., Buchanan, B.G., Jardetzky, O.: *Application of Constraint Satisfaction Techniques to the Determination of Protein Tertiary Structure*. Report KSL-86-28, Department of Computer Science, 1986.

4. Brinkley, James F., Buchanan, Bruce G., Altman, Russ B., Duncan, Bruce S., Cornelius, Craig W.: *A Heuristic Refinement Method for Spatial Constraint Satisfaction Problems*. Report KSL 87-05, Department of Computer Science.

5. Buchanan, B.G., Hayes-Roth, B., Lichtarge, O., Altman, A., Brinkley, J., Hewett, M., Cornelius, C., Duncan, B., Jardetzky, O.:*The Heuristic Refinement Method for Deriving Solution Structures of Proteins*. Report KSL-85-41. October 1985.

6. Garvey, Alan, Cornelius, Craig, and Hayes-Roth, Barbara: *Computational Costs versus Benefits of Control Reasoning*. Report KSL 87-11, Department of Computer Science.

7. Hayes-Roth, B.: *The Blackboard Architecture: A General Framework for Problem Solving?* Report HPP-83-30, Department of Computer Science, Stanford University, 1983.

8. Hayes-Roth, B.: *BB1: An Environment for Building Blackboard Systems that Control, Explain, and Learn about their own Behavior*. Report HPP-84-16, Department of Computer Science, Stanford University, 1984.

9. Hayes-Roth, B.: *A Blackboard Architecture for Control*. Artificial Intelligence 26:251-321, 1985.

10. Hayes-Roth, B. and Hewett, M.: *Learning Control Heuristics in BB1*. Report HPP-85-2, Department of Computer Science, 1985.

11. Hayes-Roth, B., Buchanan, B.G., Lichtarge, O., Hewett, M., Altman, R., Brinkley, J., Cornelius, C., Duncan, B., and Jardetzky, O.: *PROTEAN: Deriving protein structure from constraints*. Proceedings of the AAAI, 1986, p. 904-909.

12. Jardetzky, O.: *A Method for the Definition of the Solution Structure of Proteins from NMR and Other Physical Measurements: The LAC-Repressor Headpiece*. Proceedings of the International Conference on the Frontiers of Biochemistry and Molecular Biology, Alma Alta, June 17-24, 1984, October, 1984.

13. Lichtarge, Olivier: *Structure determination of proteins in solution by NMR*. Ph.D. Thesis, Stanford University, November, 1986.

14. Lichtarge, Olivier, Cornelius, Craig W., Buchanan, Bruce G., Jardetzky, Oleg: *Validation of the First Step of the Heuristic Refinement Method for the Derivation of Solution Structures of Proteins from NMR Data.*, April 1987. Submitted to Proteins: Structure, Function, and Genetics.

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 first 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.

The second goal of the project is to develop a program and set of techniques for automated summarization of patient records. The summarization program is designed to automatically create patient summaries of arbitrary and appropriate complexity as an aid for tasks such as clinical decision making, real-time patient monitoring, surveillance of quality of care, and eventually automated discovery. Two prototype summarization modules have been implemented in KEE on the Xerox 1108 workstation.

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

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.

Journal Articles

Blum, R.L.: *Computer-Assisted Design of Studies using Routine Clinical Data: Analyzing the Association of Prednisone and Cholesterol.* Annals of Internal Medicine 104(6):858-868, June, 1986.

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.

Walker, M.G. *How Feasible is Automated Discovery?* IEEE Expert 2(1):70-82, Spring, 1987.

Conference Proceedings

DeZegher-Geets, I.M., Freeman, A.G., Walker, M.G., Blum, R.L., and Wiederhold, G.C.M. *Computer-aided Summarization of a Time-oriented Medical Data Base.* In Proceedings of the Third Annual Conference on Computerization of Medical Records. Institute for Medical Record Economics, 1987.

Downs, S., Walker, M.G., and Blum, R.L. *Automated Summarization of On-line Medical Records.* In Proceedings of the Fifth World Congress on Medical Informatics (Medinfo), pages 800-804. Elsevier Science Publishers, 1986.

Walker, M.G., and Blum, R.L. *Towards Automated Discovery from Clinical Databases: the RADIX Project.* In Proceedings of the Fifth World Congress on Medical Informatics (Medinfo), pages 32-36. Elsevier Science Publishers, 1986.

National AIM Project: CADUCEUS
(INTERNIST-I)

QMR
(Quick Medical Reference)

Principal Investigators:

CADUCEUS PROJECT:
Harry E. Pople, Ph.D. (POPLE@SUMEX-AIM)
Jack D. Myers, M.D. (MYERS@SUMEX-AIM)

QMR PROJECT:
Randolph A. Miller, M.D.(RMILLER@SUMEX-AIM)
Fred E. Masarie, Jr. M.D. (MASARIE@SUMEX-AIM)
Jack D. Myers, M.D. (MYERS@SUMEX-AIM)
University of Pittsburgh
Pittsburgh, Pennsylvania 15261

Dr. Pople: (412) 624-3490
Dr. Myers: (412) 648-9933
Dr. Miller: (412) 648-3190
Dr. Masarie: (412) 648-3190

The major goal of both the CADUCEUS and INTERNIST-I/QMR Projects 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-offs as a diagnostic and triage aid to physicians' assistants, rural health clinics, military medicine and space travel. In the design of INTERNIST-I and QMR, 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 INTERNIST-1 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. QMR is broader in scope than INTERNIST-I or CADUCEUS, in that it provides quick and efficient access to the INTERNIST-I/QMR knowledge base to provide low and intermediate level informational support for physicians' decision-making, in addition to providing consultative advice.

SOFTWARE AVAILABLE ON SUMEX

Versions of INTERNIST-I 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. QMR has been shared on a restricted basis with a limited number of academic colleagues, who have agreed to give the QMR development team feedback on the program's strengths and weaknesses.

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)
Campus Box 345
Department of Psychology
University of Colorado
Boulder, Colorado 80309
(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 text analysis described by van Dijk & Kintsch (1983), yielding predictions of the recall and readability of that text by human subjects. The human-computer interaction group is developing a quantitative theory of that predicts learning, transfer, and performance for a wide range of computer-tasks, e.g. text editing, Kieras & Polson (1985).

SOFTWARE AVAILABLE ON SUMEX

A set of programs has been developed to perform the microstructure text analysis described in van Dijk & Kintsch (1983) and Kintsch & Greeno (1985). The program accepts a propositionalized text as input, and produces indices that can be used to estimate the text's recall and readability.

REFERENCES

1. Fletcher, R.C.: Understanding and solving word arithmetic problems: A computer simulation. Technical Report No. 135, Institute of Cognitive Science, Colorado, 1984.

2. Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity.* Int. J. Man-Machine Studies, 22, 365-394, 1985.

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, 1985, 92, 109-129.

5. 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, Instrumentation, & Computers, 1984, 16, 249-255.

6. Polson, P.G. and Kieras, D.E.: *A quantitative model of the learning and performance of text editing knowledge.* In Borman, L. and Curtis, B. (Eds.) Proceedings of the CHI 1985 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 207-212, 1985.

7. Polson, P.G. and Jeffries, R.: *Instruction in general problem solving skills: An analysis of four approaches*. In (Eds.) Siegel, J., Chipman, S., and Glaser, R. Thinking and learning skills: Relating instructions to basic research: Vol.1. Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 414-455.

8. Polson, P.G., Muncher, E., and Engelbeck, G.: *Test of a common elements theory of transfer*. In Mantei, M. and Orbeton, P. (Eds.) Proceedings of the CHI 1986 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 78-83, 1986.

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

10. Young, S.: A theory and simulation of macrostructure. Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

11. Walker, H.W. & Kintsch, W.: Automatic and strategic aspects of knowledge retrieval. Cognitive Science, 1985, 9, 261-283.

National AIM Project: MENTOR -- MEDICAL EVALUATION OF THERAPEUTIC ORDERS

Principal Investigators:

Stuart Speedie, Ph.D. (SPEEDIE@SUMEX-AIM)
School of Pharmacy
University of Maryland
20 N. Pine Street
Baltimore, Maryland 21201
(301) 528-7650

Terrence F. Blaschke, M.D. (BLASCHKE@SUMEX-AIM)
Department of Medicine
Division of Clinical Pharmacology
Stanford University Medical Center
Stanford, California 94305

Contact: either PI

The goal of the MENTOR project is to implement and begin evaluation of a computer-based methodology for reducing therapeutic misadventures. The project uses an on-line expert system to continuously monitor the drug therapy of individual patients and generate specific warnings of potential and/or actual unintended effects of therapy. The appropriate patient information is automatically acquired through interfaces to a hospital information system. This data is monitored by a system that is capable of employing complex chains of reasoning to evaluate therapeutic decisions and arrive at valid conclusions in the context of all information available on the patient. The results reached by the system are fed back to the responsible physicians to assist future decision making.

Specific objectives of this project include:

1. Implement a prototype computer-based expert system to continuously monitor in-patient drug therapy that uses a modular medical knowledge base and a separate inference engine to apply the knowledge to specific situations.

2. Select a small number of important and frequently occurring drug therapy problems that can lead to therapeutic misadventures and construct a comprehensive knowledge base necessary to detect these situations.

3. Design and begin implementation of an evaluation of the prototype MENTOR system with respect to its impact on the on the physicians' therapeutic decision making as well as its effects on the patient in terms of specific mortality and morbidity measures.

The work in this project builds on the extensive previous work in drug monitoring done by these investigators in the Division of Clinical Pharmacology at Stanford and the University of Maryland School of Pharmacy.

National AIM Project: SOLVER -- PROBLEM SOLVING EXPERTISE

Principal Investigators:

Paul E. Johnson, Ph.D.
School of Management and
Center for Research in Human Learning
205 Elliott Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 376-2530 (PJOHNSON@SUMEX-AIM)

James R. Slagle, Ph.D.
Department of Computer Science
136 Lind Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 373-0132 (SLAGLE@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 fifteen years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only a limited amount of expertise. 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 knowledge 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, computer hardware diagnosis, offline quality control, financial auditing, management, and law.

SOFTWARE AVAILABLE ON SUMEX

A redesigned version of the Diagnoser simulation model, named Galen, has been implemented on SUMEX. Galen is an expert system which uses recognition-based reasoning in pediatric cardiology.

REFERENCES

1. Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis*. In Bolc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

2. Johnson, P.E., Johnson, M.G., and Little, R.K.: *Expertise in trial advocacy: Some considerations for inquiry into its nature and development*, Campbell Law Review, (in press).

3. 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.

4. Johnson, P.E.: *What kind of expert should a system be?* Journal of Medicine and Philosophy, 8:77-97, 1983.

5. 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.

6. Simpson, D.E., Rich, E., Dalgaard, K., Gjerdingen, K., Crowson, T., O'Brien, D., Johnson, P.E.: *The diagnostic process in primary care: A comparison of general internists and family physicians.* SOCIAL SCIENCE AND MEDICINE (in press).

7. Spackman, K.A. and Connelly, D.P.: *Knowledge-based systems in laboratory medicine and pathology: A review and survey of the field.* Archives of Laboratory Medicine and Pathology 111:116-119, 1987.

8. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

National AIM Project: ATTENDING Project: A Critiquing Approach to Expert Computer Advice

Principal Investigator: Perry L. Miller, M.D., Ph.D.
Department of Anesthesiology
Yale University School of Medicine
New Haven, CT 06510
(203) 785-2802

Our project is exploring the "critiquing" approach to bringing computer-based advice to the practicing physician.

Critiquing is a different approach to the design of artificial intelligence based expert systems. Most medical expert systems attempt to simulate a physician's decision-making process. As a result, they have the clinical effect of trying to tell a physician what to do: how to practice medicine. In contrast, a critiquing system first asks the physician how he contemplates approaching his patient's care, and then critiques that plan. In the critique, the system discusses any risks or benefits of the proposed approach, and of any other approaches which might be preferred. It is anticipated that the critiquing approach may be particularly well suited for domains, like medicine, where decisions involve a great deal of *subjective* judgment.

To date, several prototype critiquing systems have been developed in different medical domains:

1. ATTENDING, the first system to implement the critiquing approach, critiques anesthetic management.
2. HT-ATTENDING critiques the pharmacologic management of essential hypertension.
3. VQ-ATTENDING critiques aspects of ventilator management.
4. PHEO-ATTENDING critiques the laboratory and radiologic workup of a patient for a suspected pheichromocytoma.
5. In addition, a domain-independent system, ESSENTIAL-ATTENDING, has been developed to facilitate the implementation of critiquing systems in other domains.

PUBLICATIONS

1. Miller, P.L.: Expert Critiquing Systems: Practice-Based Medical Consultation by Computer. New York: Springer-Verlag, 1986.
2. Miller, P.L. (Ed.): Selected Topics in Medical Artificial Intelligence. New York: Springer-Verlag (in press).
3. Miller, P.L., Shaw, C., Rose, J.R., Swett, H.A.: Critiquing the process of radiologic differential diagnosis. Computer Methods and Programs in Biomedicine 22:12-25, 1986.
4. Miller, P.L.: The evaluation of artificial intelligence systems in medicine. Computer Methods and Programs in Biomedicine 22:5-11, 1986.

5. Rennels, G.D., Shortliffe, E.H., Miller, P.L.: Choice of explanation in medical management: A multi-attribute model of artificial intelligence approaches. Medical Decision Making 7:22-31, 1987.

6. Mars, N.J.I., Miller, P.L.: Knowledge acquisition and verification tools for medical expert systems. Medical Decision Making 7:6-11, 1987.

7. Miller, P.L., Blumenfrucht, S.J., Rose, J.R., Rothschild, M., Swett, H.A., Weltin, G., Mars, N.J.I.: HYDRA: A knowledge acquisition tool for expert systems which critique medical workup. Medical Decision Making 7:12-21, 1987.

8. Swett, H.A., Miller, P.L.: ICON: A computer-based approach to differential diagnosis in radiology. Radiology (in press).

9. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: A computational model of reasoning from the clinical literature. Computer Methods and Programs in Biomedicine (in press).

10. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: A structured representation of the clinical literature and its use in a medical management advice system. Bulletin du Cancer (in press).

11. Miller, P.L.: Exploring the critiquing approach: Sophisticated practice-based feedback by computer. Proceedings of the Fifth World Conference on Medical Informatics: MEDINFO-86, Washington, D.C., October 1986, pp 2-6.

12. Mars, N.J.I., Miller, P.L.: Tools for knowledge acquisition and verification in medicine. Proceedings of the Tenth Symposium on Computer Applications in Medical Care, Washington, D.C., October 1986, pp. 36-42.

13. Miller, P.L., Blumenfrucht, S.J., Rose, J.R., Rothschild, M., Weltin, G., Swett, H.A., Mars, N.J.I.: Expert system knowledge acquisition for domains of medical workup: An augmented transition network model. Proceedings of the Tenth Symposium on Computer Applications in Medical Care, Washington, D.C., October 1986, pp. 30-35.

14. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: Reasoning from the clinical literature: The Roundsman system. Proceedings of the Fifth World Conference on Medical Informatics: MEDINFO-86, Washington, D.C., October 1986, pp. 771-775.

15. Rennels, G.D., Shortliffe, E.H., Stockdale, F.E., Miller, P.L.: Updating an expert knowledge base as medical knowledge evolves: Examples from oncology management. Proceedings of the American Association of Medical Systems and Informatics Congress 87, San Francisco, May 1987, pp. 228-231.

16. Fisher, P.R., Miller, P.L., Swett, H.A.: A script-based representation of medical knowledge involving multiple perspectives. Proceedings of the American Association of Medical Systems and Informatics Congress-87, San Francisco, May 1987, pp. 233-237.

17. Miller, P.L.: Expert consultation systems in medicine: A complex and fascinating domain. Proceedings of the Annual Meeting of the IEEE (Electro-87), New York, April 1987, pp. 1/2:1-4 (invited paper).

Stanford Project: REFEREE Project

Principal Investigators: Bruce G. Buchanan, Principal Investigator
Computer Science Department
Stanford University
Stanford, California 94305

Byron W. Brown, Co-Principal Investigator
Department of Medicine
Stanford University Medical Center
Stanford, California 94305

Daniel E. Feldman, Associate Investigator
Department of Medicine
Stanford University Medical Center
Stanford, California 94305

The goals of this project are related both to medical science and Artificial Intelligence: (a) use AI methods to allow the informed but non-expert reader of the medical literature to evaluate a randomized clinical trial, and (b) use the interpretation of the medical literature as a test problem for studies of knowledge acquisition and fusion of information from disparate sources. REFEREE and REVIEWER, a planned extension, will be used to evaluate the medical literature of clinical trials to determine the quality of a clinical trial, make judgements on the efficacy of the treatment proposed, and synthesize rules of clinical practice. The research is an initial step toward a more general goal - building computer systems to help the clinician and medical scientist read the medical literature more critically and more rapidly.

National AIM Project: Computer-Aided Diagnosis of Lymph Node Pathology (PATHFINDER)

Principal Investigator: Bharat Nathwani, M.D.
Department of Pathology
HMR 204
2025 Zonal Avenue
University of Southern California
School of Medicine
Los Angeles, California 90033
(213) 226-7064 (NATHWANI@SUMEX-AIM)

Lawrence M. Fagan, M.D., Ph.D.
Medical Computer Science Group
Department of Medicine
Medical School Office Building
Stanford, California 94305
(415) 723-6979 (FAGAN@SUMEX-AIM)

The PATHFINDER Project is centered on the construction of an expert system for assisting pathologists with the diagnosis of tissue pathology. PATHFINDER research is focused on the domain of lymph node pathology. The project is based at the University of Southern California in collaboration with the Stanford University Medical Computer Science Group. Ongoing AIM research has been addressing fundamental problems of knowledge representation, reasoning strategies, user modeling, explanation, and user acceptance. A pragmatic goal of the project is to provide a valuable diagnostic and educational tool for pathologists with different levels of training and experience by integrating diverse knowledge about lymph node pathology. It is hoped that PATHFINDER basic research on representation and inference in combination with the pragmatic goals of constructing a clinically-relevant diagnostic aid will lead to useful advances in medical computing.

A pilot version of the program provides diagnostic advice on eighty common benign and malignant diseases of the lymph nodes based on 150 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.

AIM Pilot Project: RXDX Project

Principal Investigators:
Robert Lindsay, Ph.D. (313) 764-4227
Michael Feinberg, M.D., Ph.D. (215) 842-4208
University of Michigan
Ann Arbor, Michigan

We are developing a prototype expert system that could 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 program uses a knowledge base constructed by encoding the clinical expertise of a skilled psychiatrist in a set of rules and other knowledge structures. 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).

The potential benefits to psychiatry include: making relatively skilled psychiatric consultation widely available in underserved areas, including some public mental health facilities where patients are seen by non-psychiatrists and have relatively little direct patient-physician contact; providing non-psychiatrically trained physicians with additional information about psychiatric diagnosis and treatment; avoiding errors of oversight caused by inaccessible patient data; and increased productivity in patient care. Like any good consultant, the program will be able to teach the interested user, and can function as a teaching tool independent of direct clinical application.

PUBLICATIONS

1. Feinberg, M. and Lindsay, R. K.: *Expert systems in Psychiatry and Psychopharmacology*. Psychopharmacol. Bull., 22, 1986, 311-316.

2. Lindsay, R. K. Expert Systems in Psychiatric Diagnosis: Rule-Based Systems. Presented at MedInfo86, Washington, D.C.

3. Feinberg, M. What Psychiatrists Can't Do. Presented at Medinfo86, Washington, D. C.

National AIM Project: DECISION SUPPORT FOR TIME-VARYING CLINICAL PROBLEMS

Principal Investigator: Lawrence Widman, M.D., Ph.D.
Division of Cardiology
Case Western Reserve University
2065 Adelbert Road
Cleveland, OH 44106
(216) 844-3153

Time-varying systems, which include many areas of medicine, science, economics, and business, can be described mathematically by differential equations. They are distinct from the pattern-matching and logic-based domains dealt with so successfully by existing expert system methods, because they can include feedback relationships. It is generally felt that they are best approached by enhancement of existing methods for deep model-based reasoning.

The goal of this project is to develop AI methods for capturing and using knowledge about time-varying systems. The strategy is to address general problems in model-based knowledge representation and reasoning. The intermediate objective is to develop methods which are powerful enough to work in selected realistic situations yet are general enough to be transportable to other, unrelated knowledge domains.

The tactical approach is to work on well-defined yet complex and interesting problems in the medical domain. We have, therefore, selected the human cardiovascular system as our prototype of a time-varying system, and are developing methods for representing and reasoning about its mechanical and electrical activities in the normal and diseased states.

REFERENCES

1. Widman, L.E. Reasoning about Diagnosis and Treatment in a Causal Medical Model using Semi-Quantitative Simulation and Inference. Workshop on Artificial Intelligence in Medicine, National Conference on Artificial Intelligence, AAAI-87, Seattle.

2. Widman, L.E., Lee, Y.-B., and Y.-H. Pao. Diagnosis of Causal Models by Semi-Quantitative Reasoning (submitted to SCAMC 1987).

3. Widman, L.E., Lee, Y.-B., and Y.-H. Pao. Diagnosis of Causal Medical Models by Semi-Quantitative Reasoning. In: Miller, P.L. (ed.) Topics in Medical Artificial Intelligence, Springer-Verlag (in preparation).

4. Lee, Y.-B. and L.E. Widman. Reasoning about Diagnosis and Treatment in a Causal Time-varying Domain using Semi-Quantitative Simulation and Inference. Workshop on Artificial Intelligence and Simulation, National Conference on Artificial Intelligence, AAAI-86, Philadelphia.

5. Widman, L.E. Representation Method for Dynamic Causal Knowledge Using Semi-Quantitative Simulation. Fifth World Conference on Medical Informatics. 1986: 180-184.

National AIM Project: KNOWLEDGE ENGINEERING FOR RADIATION THERAPY

Principal Investigator: Ira J. Kalet, Ph.D.
School of Medicine
University of Washington at Seattle
Seattle, Washington 98195
(206) 548-4107

We are developing an expert system for planning of radiation therapy for head and neck cancers. The project will ultimately combine knowledge-based planning with numerical simulation of the radiation treatments. The numerical simulation is needed in order to determine if the proposed treatment will conform to the goals of the plan (required tumor dose, limiting dose to critical organs). The space of possible radiation treatments is numerically very large, making traditional search techniques impractical. Yet, with modern radiation therapy equipment, the design of treatment plans might be significantly aided by automatically generating plans that meet the treatment constraints. The project will result in systematization of knowledge about radiation treatment design, and will also provide an example of how to represent and solve design problems with a knowledge based system.

This project has some relevance to computer science as well, in that our approach, if successful, may contribute to a better understanding of design problem solving with knowledge-based systems.

REFERENCES

I. Kalet and W. Paluszynski: A Production Expert System for Radiation Therapy Planning. Proceedings of the AAMSI Congress 1985, May 20-22, 1985, San Francisco, California. Edited by Allan H. Levy and Ben T. Williams. American Association for Medical Systems and Informatics, Washington, D.C., 1985.

W. Paluszynski and I. Kalet: Radiation Therapy Planning: A Design Oriented Expert System. WESTEX-87 (Western Conference on Expert Systems), Anaheim, California, June 2-4, 1987.

I. Kalet and J. Jacky: Knowledge-based Computer Simulation for Radiation Therapy Planning. Proceedings of the Ninth International Conference on the use of Computers in Radiotherapy, Scheveningen, the Netherlands, June 1987. North Holland, 1987.

AIM Pilot Project: COMPUTER-BASED EXERCISES IN PATHOPHYSIOLOGIC DIAGNOSIS

Principal Investigator: J. Robert Beck, M.D.
MIS Butler I
2 Maynard St.
Dartmouth College
School of Medicine
Hanover, NH 03755
(603) 646-7171

Research in artificial intelligence at Dartmouth Medical School focuses on three main areas: 1) knowledge-based systems applied to laboratory medicine and pathology, 2) knowledge acquisition using machine learning techniques, and 3) computer-based instruction using artificial intelligence techniques to critique students' workup plans. These projects have in common the fundamental research questions of how knowledge should be represented and used in a classification approach to problem-solving related to the use of laboratory data.

An interdisciplinary team of computer scientists, physicians, and educators is working on the Computer-based Exercises project. A prototype system is nearing completion, with formative evaluation scheduled for Fall, 1987.

REFERENCES

Beck, J.R., Prietula, M.J., Russo, E.A.: A role for intelligent systems in teaching medical pathophysiology. In: Salamon, R., Blum, B., Jorgenson, M. (eds). Proc. Fifth Conf. Med. Inform. (MEDINFO '86), Elsevier-North Holland, Amsterdam, 1986, 936-938.

Beck, J.R.: Artificial intelligence: A topic for Medical Decision Making? (edit.) Med. Decis. Making 1987; 7:4.

References

1. Barr, Avron, Paul R. Cohen and Edward A. Feigenbaum. *The Handbook of Artificial Intelligence, Volumes I, II, and III.* William Kaufmann, Inc., Los Altos, CA, 1981 and 1982.

2. Congressional Research Service, Library of Congress. The Impact of Information Technology on Science. Science Policy Study Background Report Number 5, Task Force on Science Policy, Committee on Science and Technology, U.S. House of Representatives, September, 1986.

3. Coulter, C. L. "Research Instrument Sharing." *Science 201*, 4354 (1978).

4. Harmon, Paul and King, David. *Expert Systems - Artificial Intelligence in Business*. John Wiley & Sons, Inc., New York, NY, 1985.

5. Hayes-Roth, F., Waterman, D. A. & Lenat, D. B. (Eds.). *Building Expert Systems*. Addison-Wesley, Reading, MA, 1983.

6. Horvitz, E.J., Heckerman, D.E., and Langlotz, C.P. A Framework for Comparing Alternative Formalisms for Plausible Reasoning. Proceedings of AAAI-86, Fifth National Conference on Artificial Intelligence, American Association for Artificial Intelligence, Menlo Park, CA, August, 1986, pp. 210-214.

7. Information Technology R&D Project Staff. Information Technology R&D: Critical Trends and Issues (Summary). Report OTA-CIT-269, Congressional Office of Technology Assessment, October, 1985.

8. Keith A. Lantz et. al. *V - System 4.1 Reference Manual.* December 1, 1983. Computer Systems Laboratory, Stanford University.

9. Keith A. Lantz, David R. Cheriton, and William I. Nowicki. Third Generation Graphics for Distributed Systems. Computer Systems Laboratory. Stanford University, December 21, 1982

10. Lederberg, J. "Digital Communications and the Conduct of Science: The New Literacy." *Proceedings of the IEEE 66*, 11 (1978).

11. Nilsson, N.. *Principles of Artificial Intelligence*. Tioga, Palo Alto, CA, 1980.

12. NLM Planning Panel 4. NLM Long Range Plan: Medical Informatics. Shortliffe, E.H., Panel 4 Chairman , National Library of Medicine, January, 1987.

13. Rich, E.. *Artificial Intelligence*. McGraw-Hill, New York, 1983.

14. Paul S. Rosenbloom, John E. Laird. Mapping Explanation-Based Generalization onto SOAR. Proceedings of the Fifth National Conference on Artificial Intelligence, American Association for Artificial Intelligence, Philadelphia, Penn., August 11-15, 1986, pp. 561-567. Also KSL-86-46 and STAN-CS-86-1111

15. Scheifler, R.W. and Gettys, J. The X Window System. Draft October 1986. To appear in a special issue of the *ACM Transactions on Graphics* -- user interface software

16. Steele, Guy L., Jr.. *Common Lisp - The Language*. Digital Press, Burlington, MA, 1984.

17. *NeWS Preliminary Technical Overview*. Sun Microsystems, Inc., 1986.

18. Waterman, Donald A.. *A Guide to Expert Systems*. Addison-Wesley Publishing Company, Menlo Park, CA, 1986.

19. David C. Wilkins and Bruce G. Buchanan. On Debugging Rule Sets When Reasoning Under Uncertainty. Tech. Rept. KSL 86-30, Stanford University, Knowledge Systems Laboratory, July, 1986. Earlier version appeared in Proceedings of the Fifth National Conference on Artificial Intelligence, August 1986

20. David C. Wilkins, Willian J. Clancey and Bruce G. Buchanan. Using and Evaluating Differential Modeling in Intelligent Tutoring and Apprentice Learning Systems. Tech. Rept. KSL 86-62, Stanford University, Knowledge Systems Laboratory, October, 1986.

21. David C. Wilkins. Knowledge Base Debugging Using Apprenticeship Learning Techniques. Tech. Rept. KSL 86-63, Stanford University, Knowledge Systems Laboratory, December, 1986.

22. David C. Wilkins, William J. Clancey and Bruce G. Buchanan. Knowledge Base Refinement Using Abstract Control Knowledge. Tech. Rept. KSL 87-01, Stanford University, Knowledge Systems Laboratory, January, 1987.

23. Winston, P.. *Artificial Intelligence, 2nd ed.* Addison-Wesley, Reading, MA, 1984.