SUMEX

STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE

RR-00785

ANNUAL REPORT—YEAR 12

Submitted to

BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM
NATIONAL INSTITUTES OF HEALTH

June 1, 1985

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 RR00785-12
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:	08-01-84
5b. TO:	07-31-85
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:	July 5, 1985
8. TELEPHONE:	415-497-6979

Table of Contents

List of Figures

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 18, our training activities are summarized starting on page 49, and the progress of our collaborating projects is detailed starting on page 85.

II.B. Books, Papers, and Abstracts

The list of recent publications for our core research and development work starts on page 45 and those for the collaborating projects are in the individual reports starting on page 85.

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 52.

III. Narrative Description

We are about to start the final grant year in the current SUMEX-AIM award. This annual report was prepared in parallel with a competing renewal proposal for continuing the resource beyond July 1986. The report is based on the comprehensive progress sections of the proposal for resource core research and development and for the collaborating scientific community.

III.A. Summary of Research Progress

III.A.1. Executive Summary

This summary provides an overview of our accomplishments. In the almost twelve years since the SUMEX-AIM resource was established, computing technology and biomedical artificial intelligence research have undergone a remarkable evolution. As we prepare to renew the resource through the remainder of the 1980's, we take pride in the realization that SUMEX has both influenced and responded to those changing technologies. It is widely recognized that our resource has fostered highly influential work in medical AI -- work from which it is generally acknowledged that the expert systems field emerged -- and that it has simultaneously helped define the technological base of applied AI research. The LISP machines 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. We look forward to another half decade during which the era of centralized machines for AI research will come to an end, having been supplanted by networks of distributed and heterogeneous single-user machines sharing common resources such as file servers, printers, and gateways to other local and long-distance networks.

We continue to be motivated by three main goals:

1. to develop and provide impeccable computing resources and human assistance to scientists working on applications of artificial intelligence research in medicine and biology;

2. to demonstrate that it is feasible to provide resources and assistance to a national community of researchers from a central site, integrating distributed and centralized computing technology, local and national computer communication networks, and a staff oriented toward the special problems of individuals participating in AIM research at other institutions;

3. to develop the community of scientists interested in working on applications of AI to the biomedical sciences; facilitating the growth, health, and vigor of the community by providing electronic communications that link its members and by assisting with the dissemination of systems software and applications programs that are of use to the wider community of AIM researchers. One question we have been asking is, "Is there a new style of science that will emerge in a communications-enhanced setting of national, rather than institutional, scope?" Within a decade it was clear that the answer to this question was (and is) "yes"!

SUMEX's Success as a National Research Resource

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. Many NIH computer RR's have been mostly institutional in scope, occasionally regional (like the UCLA resource). SUMEX now has the reputation of 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.

A decade ago, when machines up to the task of supporting AI research cost $1M, some of the most notable projects in the history of Artificial Intelligence were done with terminal-and-network, without a computer on site. In human terms, this meant, of course, not having the headaches and energy drains of proposing a machine, installing it, maintaining it and its software, hiring its system programmers and operators, dealing with communication vendors, etc. The famous INTERNIST program was developed from Pittsburgh in this way. And the ACT computer model was begun at Michigan, continued at Yale, and later at Carnegie-Mellon, all without moving the program or losing a day's work because of machine transition problems. The GENET community of over 300 molecular biologists grew up in a year around SUMEX programs for analyzing DNA sequences. Their demand for these centralized capabilities ultimately swamped our machine and led to the initiation of a separate resource (BIONET) to meet their needs.

The projects SUMEX supports have generally required substantial computing resources with excellent interaction. Even today though, with the growing availability of Lisp 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. For example, computer scientists have been collaborating fruitfully with physical chemists, molecular biochemists, geneticists, crystallographers, internists, ophthalmologists, infectious disease specialists, intensive care specialists, oncologists, psychologists, biomedical engineers, and other expert practitioners. And in some of these cases, the interdisciplinary collaboration, usually so difficult to achieve in the best of circumstances, was achieved in spite of geographical distance between the participants, using the computer networks.

SUMEX has also achieved successes as a community builder. 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 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 IntelliCorp; the Rutgers Computers in Biomedicine resource is centered at Rutgers University; the CADUCEUS project splits their research work between their own VAX computer and the SUMEX resource; and the Chemical Synthesis project now operates entirely on a VAX at U.C. Santa Cruz.

The SUMEX mission has been able to capture the contributions of some of the finest computers-in-medicine specialists and computer scientists in the country. For example, Professor Joshua Lederberg (SUMEX's first PI, now President of The Rockefeller University) is a member of SUMEX's Executive Committee; and Dr. Donald Lindberg, former Director of the University of Missouri's Medical Information Science group, and now Head of the National Library of Medicine, was until recently the Chairman of the AIM Advisory Group. Professor Herbert Simon of Carnegie-Mellon University, Professor Marvin Minsky of MIT, and many other distinguished scientists serve on that peer review committee.

SUMEX and Artificial Intelligence Research

The SUMEX Project is a relative latecomer to AI research. Yet its scope has given strong impetus to this historic development in applied computer science. AI research is that part of computer science that investigates symbolic reasoning processes, and the representation of symbolic knowledge for use in inference. It views heuristic or judgmental knowledge to be of equal importance with "factual" knowledge, indeed to be the essence of what we call "expertise". In its "Expert Systems" work, it seeks to capture the expertise of a field, and translate it into programs that will offer intelligent assistance to a practitioner in that field.

For computer applications in medicine and biology, this research path is crucial, indeed ineluctable. 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.

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. The AI methodology is beginning to fill that need. Heuristic knowledge can be elicited, studied, critiqued by peers, and taught to students.

The tide of AI research and application is rising. AI is one of the principal fronts along which university computer science groups are expanding. Federal and industrial support for AI research is vigorous and growing, although support specifically for biomedical applications continues to be limited. 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% 4 years ago). 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 our 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.

There is 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 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.

In industry, AI is on an exponential growth path as well. In the USA alone, over 30 AI start-up companies have been formed in the past four years and many groups have been established in large companies as well. The list of names is long and includes Hewlett-Packard, Schlumberger (including Fairchild), Texas Instruments, Xerox, IBM, DEC, General Motors, General Electric, Boeing, Rockwell, FMC Corp, Ford-Aerospace, Apple Computer, Teknowledge, IntelliCorp, Syntelligence, Lucid, Inference Corp, Symbolics, LMI, and so on... Many of these firms are marketing hardware and software tools for expert system development, as well as custom system services. And Japan has mounted a long-term, well-funded "Fifth Generation" computing effort to broadly develop knowledge-based systems technology as part of their national economic base of the 1990's.

The AI tide is rising 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. First among these scientific directions has been a switch in AI's focus from inference procedures to knowledge representation and use. There is now a recognition that the power of problem solvers derives primarily from the knowledge that they contain -- of the elements of the problem domain, of the strategies for solving problems in that domain, and of the forms in which the knowledge is to be acquired. In 1977, Goldstein and Papert of MIT, writing in the journal Cognitive Science, described the change of focus as a "paradigm shift" in AI. This shift was induced largely (though, of course, not exclusively) by the work at SUMEX, beginning with the DENDRAL development in 1965.

Toward the '90s: the Future of SUMEX

Given this setting of success and vitality, 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, with the emergence of powerful single-user workstations and local area networking? The answers remain clear.

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. Our growing knowledge is clearly attained in an incremental fashion; we build today on the results of the past decade, and we will build in the 1990's on the work we undertake today.

At the resource level, there is a growing, diverse, and active AIM research community with intense needs for computing resources to continue its work. Many of these groups still are dependent on the SUMEX-AIM resources. For those who 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 (previously called the Heuristic Programming Project), make this an ideal setting to integrate, experiment with, and export these methodologies for the rest of the AIM community.

III.A.2. Resource Goals and Definitions

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. What is Artificial Intelligence?

Artificial Intelligence research is that part of Computer Science concerned with symbol manipulation processes that produce intelligent action [1, 26, 29, 35]. Here *intelligent action* means an act or decision that is goal-oriented, is arrived at by an understandable chain of symbolic analysis and reasoning steps, and utilizes knowledge of the world to inform and guide the reasoning.

Placing AI in Computer Science

A simplified view relates AI research with the rest of computer science. The manner of use of computers by people to accomplish tasks can be thought of as a one-dimensional spectrum representing the nature of the instructions that must be given the computer to do its job. At one extreme of the spectrum, representing early computer science, the user supplies his intelligence to instruct the machine precisely *how* to do the job, step-by-step.

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

Progress in computer science may be seen as steps away from that extreme *how* point on the spectrum: the familiar panoply of assembly languages, subroutine libraries, compilers, extensible languages, etc. illustrate this trend. The research activity aimed at creating computer programs that act as *intelligent agents* near the *what* end of the spectrum can be viewed as a long-range goal of AI research.

Expert Systems and Applications

The national SUMEX-AIM resource has in large part made possible a long, interdisciplinary line of artificial intelligence research at Stanford concerned with the development of concepts and techniques for building expert systems [15]. An *expert* *system* is an intelligent computer program that uses knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution. For some fields of work, the knowledge necessary to perform at such a level, plus the inference procedures used, can be thought of as a model of the expertise of the expert practitioners of that field.

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

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 C on page 215 and more detailed progress summaries in Section IV on page 85. 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. There have already been promising results in areas such as chemical structure elucidation and synthesis, diagnostic consultation, molecular biology, and modeling of psychological processes.

Needless to say, much is yet to be learned in the process of fashioning a coherent scientific discipline out of the assemblage of personal intuitions, mathematical procedures, and emerging theoretical structure comprising artificial intelligence research. State-of-the-art programs are far more narrowly specialized and inflexible than the corresponding aspects of human intelligence they emulate; however, in special domains they may be of comparable or greater power, e.g., in the solution of structure problems in organic chemistry or in the rigorous consideration of a large diagnostic knowledge base.

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 [19, 5]. And, as projected earlier, 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 highly desirable latitude 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. As planned, this central role has already evolved significantly and continues to evolve 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 actively supported efforts by the more mature AIM projects to develop or adapt additional computing facilities tailored to their particular needs and designed to free the main SUMEX resource for new, developing applications projects. To date, over 10 of the national projects have moved some or all of their work to local sites and several have begun resource communities of their own (see page 79). 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 to Biomedicine

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

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

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

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

We believe that computers are an inevitable technology for helping to overcome these problems. While recognizing the value of mathematical modeling, statistical classification, decision theory and other techniques, we believe that effective use of such methods depends on using them in conjunction with less formal knowledge, including contextual and strategic knowledge.

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

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

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

Evidence of the impact of SUMEX-AIM in promoting ideas such as these, and developing the pertinent specific techniques, has been the explosion of interest in medical artificial intelligence and the specific research efforts of the SUMEX community. In SUMEX's second decade, we have found that the small community of researchers that characterized the AIM field in the early 1970's has now grown to a large, accomplished, and respected research community. The American Association for Artificial Intelligence (AAAI), the principal scientific membership organization for the AI field, has 7000 members, over 1000 of whom are members of the medical special interest group known as the AAAI-M. This subgroup was founded by members of the SUMEX-AIM community who were active in AAAI and is the only active subgroup in the Association. The organization distributes semiannual newsletters on medical AI and provides a focus for cosponsoring relevant medical computing meetings with other societies (such as the American Association for Medical Systems and Informatics -- AAMSI). Medical AI papers are prominently featured at both medical computing and artificial intelligence meetings, and artificial intelligence is now routinely featured as a specific subtopic for specialized sessions at medical computing and other medical professional meetings. For example, members of the AIM community have represented the field to physicians at the American College of Pathology and American College of Physicians meetings for the last several years. A mere decade ago, the words "artificial intelligence" were never uttered at such conferences. The growing interest and recognition are largely due to the activities of the SUMEX-AIM community.

Another indication of the growing impact of the SUMEX-AIM community is its effect on medical education. For reasons such as those outlined above, there is an increasing recognition of the need for a revolution in the way medicine is taught and medical students organize and access information. Computing technology is routinely cited as part of this revolution, and artificial intelligence (and SUMEX-AIM research) generally figures prominently in such discussions. Such diverse organizations as the National Library of Medicine, the American College of Physicians, the Association of American Medical Colleges, and the Medical Library Association have all called for sweeping changes in medical education, increased educational use of computing technology, enhanced research in medical computer science, and career development for people working at the interface between medicine and computing; reports of all four organizations have specifically cited the role of artificial intelligence techniques in future medical practice and have used SUMEX-AIM programs as examples of where the technology is gradually heading.

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

III.A.2.4. Summary of Current Goals

The following summarizes SUMEX-AIM resource objectives as stated in the proposal for the on-going five-year grant, begun on August 1, 1981, and provides the backdrop against which specific progress is reported. These project goals are presented in the three categories used in the previous proposal: 1) resource operations, 2) training and education, and 3) core research.

1) Resource Operations

- Maintain the vitality of the AIM community by continuing to encourage and explore new applications of AI to biomedical research and improving mechanisms for inter- and intra-group collaborations and communications. User projects will fund their own manpower and local needs; will actively contribute their special expertise to the SUMEX-AIM community; and will receive an allocation of computing resources under the control of the AIM management committees. There will be no "fee for service" charges for community members.

- Provide effective computational support for AIM community goals, including efforts to improve the support for artificial intelligence research and new applications work; to develop new computational tools to support more mature projects; and to facilitate testing and research dissemination of nearly operational programs. We will continue to operate and develop the existing KI-10/2020 facility as the nucleus of the resource. We will acquire additional equipment to meet developing community needs for more capacity, larger program address spaces, and improved interactive facilities. New computing hardware technologies becoming available now and in the next few years will play a key role in these developments and we expect to take the lead in this community for adapting these new tools to biomedical AI needs. We planned the phased purchase of two VAX computers to provide increased computing capacity and to support large address space LISP development, a 2 GByte file server to meet file storage needs, and a number of single-user "professional workstations" to experiment with improved human interfaces and AI program dissemination.

- 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 TYMNET connections for at least the near term and will actively explore other advantageous connections to new communications networks and to dedicated links.

2) Training and Education

- Provide community-wide support and work to make resource goals and AI programs known and available to appropriate medical scientists. Collaborating projects are responsible for the development and dissemination of their own AI programs.

- Provide documentation and assistance to interface users to resource facilities and programs and continue to exploit particular areas of expertise within the community for developing pilot efforts in new application areas.

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

- Support workshop activities, including collaboration with the Rutgers Computers in Biomedicine resource on the AIM community workshop and with individual projects for more specialized workshops covering specific application areas or program dissemination.

3) Core Research

- Explore basic artificial intelligence research issues and techniques, including knowledge acquisition, representation, and utilization; reasoning in the presence of uncertainty; strategy planning; and explanations of reasoning pathways, with particular emphasis on biomedical applications.
- Support community efforts to organize and generalize AI tools that have been developed in the context of individual application projects. This will include work to organize the present state-of-the-art in AI techniques through the AI Handbook effort and the development of practical software packages (e.g., AGE, EMYCIN, UNITS, and EXPERT) for the acquisition, representation, and utilization of knowledge in AI programs.

III.A.3. Details of Technical Progress

This progress summary covers only the resource nucleus. 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 4 years, focusing on the resource nucleus.

- 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. Nine projects now operate on their own facilities, including three that have become BRTP resources in their own right. Nine projects in the community have completed their research goals and their staffs have moved on to new areas.

- SUMEX user projects have made good progress in developing and disseminating effective consultative computer programs for biomedical research. These performance programs provide expertise in analytical biochemical analyses and syntheses, clinical diagnosis and decision-making, molecular biology, and various kinds of cognitive and affective psychological modeling. We have worked hard to meet their needs and are grateful for their expressed appreciation (see Section IV).

- We have made significant strategic improvements to the SUMEX-AIM computing environment in order to optimize computing support for the community. These developed in ways somewhat different from the initially projected plan. The DEC VAX computer did not prove to be an effective machine for running Lisp [23], while Lisp workstations have in fact become available from a number of vendors as tentatively expected at the time of our proposal (first Xerox, then Symbolics and LMI, and more recently Hewlett-Packard and Texas Instruments). Thus, rather than augmenting our mainframe resources with the purchase of large address space VAX's, we upgraded the KI-TENEX system to a DEC 2060 and at the same time, began moving aggressively toward a Lisp workstation-based research environment, with the approval of an ad hoc site visit group. We did secure VAX capabilities for our community by means of access to an 11/780 purchased under DARPA funding. We made an initial purchase of Xerox Dolphins with NIH funding and subsequently added more Xerox and Symbolics machines with NIH and DARPA funding and with industrial gifts. Because of the broad mix of research in the SUMEX-AIM community, no single workstation 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 the dissemination of SUMEX-AIM technology through various media. We have distributed various AI software tools to many research laboratories, including over 200 combined copies of the GENET, EMYCIN, AGE, MRS, SACON, GLISP, and BB-1 systems. Several of our software systems have been adapted as commercial AI tools such as the Teknowledge S.1 and M.1 systems derived from EMYCIN, the Texas Instruments Personal Consultant system derived from EMYCIN, and the IntelliCorp KEE system derived from UNITS. We have also prepared video tapes of some of our research projects including ONCOCIN and an overview tape of Knowledge Systems Laboratory work.

- 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 (see page 81). These books have included the three-volume set of the *Handbook of Artificial Intelligence*, edited by Barr, Cohen, and Feigenbaum; a book on *Readings in Medical Artificial Intelligence: The First Decade* by Clancey and Shortliffe; and a book on *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project* by Buchanan and Shortliffe.

- We completed the GENET project, begun in 1980 as a collaboration between the MOLGEN investigators and SUMEX, to make a set of DNA sequence analysis computing tools available to a national community of molecular biologists. This was an experiment in using a SUMEX-like resource to disseminate sophisticated software tools to a computer-naive community and proved extremely successful. GENET served over 300 molecular biologists before being phased out in early 1983. Subsequently, a new resource called BIONET has been funded by NIH at IntelliCorp to provide routine service of the type pioneered by SUMEX/GENET.

- A program in Medical Information Sciences was begun at Stanford in 1983 under Professor Shortliffe as Director. 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 focusses 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 feasibility of this program resulted in large part from the prior work and research computing environment provided by the SUMEX-AIM resource. Over 20 PhD and MS trainees will be enrolled in the fall of 1985. It has been awarded post-doctoral training support from the National Library of Medicine, received an equipment gift from Hewlett-Packard, and has received additional industrial and foundation grants for student support.

- We made significant progress in core AI research. In the area of knowledge representation, work was done on the representation of explicit strategy knowledge, temporal knowledge, causal knowledge, and knowledge in logic-based systems. In the area of architectures and control, we worked on a new implementation of a blackboard architecture with explicit control knowledge. Under knowledge acquisition studies, three PhD theses were completed covering experiments in learning by induction, by analogy, and learning from partial theories. In the area of knowledge utilization, results include work on reasoning with uncertainty and using counterfactual conditionals. We continued work on a number of existing tools for expert systems and on building new ones such as the BB1 system. And finally, significant work was done on the inference of user models, skeletal planning, defining a taxonomy of diagnostic methods, and reasoning with causal models.

- We have continued the core development of the SUMEX facility hardware, software, and 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 machines based on MC-68000 microprocessors were developed to link our environment together and are now the standard system used in the campus-wide Stanford University network. We developed a gateway interface between Apple equipment (e.g., the Macintosh and Lisa) and EtherNet hosts that is now in wide use at universities around the country. We have developed many other software packages to enhance the computing environments of the Lisp workstations and to link them to other hosts and servers on our networks.

III.A.3.2. Resource Equipment Details

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 which was extended locally to support a dual processor configuration. Because of the operational load on the KI-10's, in the late 1970's, we had added a small DEC 2020 system (see Figure 2) to support more dedicated testing of systems like ONCOCIN and Caduceus and for community demos. This facility provided a superb base for the AI mission of SUMEX-AIM through 1982. Its interactive computing environment, its AI program development tools, and its network and interpersonal communication media were 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 also give us access to a large computer science research community, including active artificial intelligence and system development research groups.

The Heterogeneous Computing Environment

In the renewal for the current grant period, both an augmentation of the central resource in terms of address space and capacity and exploratory work with Lisp workstations were planned. The Initial Review Group recognized in their special study section report the importance of optimizing the timing of our planned hardware acquisitions to coordinate community needs with the availability of important technological developments in vendor-supported systems. They recommended in their report that we be allowed considerable flexibility as to phasing of equipment purchases within the 5-year renewal period.

We had initially planned to purchase a large VAX in 1981 and later, our first Lisp workstations. However, we speeded our push toward workstations for several reasons. The state of VAX Lisp implementations and projections of their performance were very discouraging (a study of the VAX InterLisp implementation was done at the time as documented in [23]). And the first Xerox InterLisp Dolphin workstations were available for delivery after the summer of 1981. These machines were the prototypes on which research toward adapting expert AI systems for the interactive workstation environment could begin. So, we purchased 5 Dolphins for the fall of 1981 and, in order not to delay non-Lisp SUMEX-AIM work involving VAX machines, we were able to arrange shared access to a VAX 11/780 funded by ARPA to support Heuristic Programming Project research. One of the Dolphins we purchased was loaned for several years to the Rutgers Computers in Biomedicine resource for experimental work.

We continued to evaluate strategies and alternatives for planned system configuration development. In particular, we had a chance to gain experience with the Dolphin InterLisp machines and the shared VAX, reassess the role of the dual KI-TENEX system, and reach a consensus about what the long term configuration of the SUMEX-AIM facility should be. This was validated by an ad hoc study section review in 1982. In summary, it was decided that the best resource configuration for the coming decade would be a shared central machine coupled through a high-performance network to growing clusters of personal workstations. The central machine should be an extended addressing TOPS-20 machine and the workstations will be chosen from the viable products available and scheduled for announcement.

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 these systems were their processing speed and cost, but the prospect of other workstations to be available from Xerox, Symbolics, LMI, HP, and others reassured us that these were the right choices for AI system in the long term. Still, at the time, it was not possible to equip very much of the SUMEX-AIM community with individual workstations.

Upgrade of the KI-10's to a 2060

Meanwhile, on the mainframe front, given the continued need for a central machine, the poor Lisp performance of the VAX, and the increasingly untenable difficulties in maintaining the KI-TENEX system, we decided it is time to retire the KI-10's and upgrade them to the then (1982) more modern DEC 2060 TOPS-20 system. This would free our systems staff to concentrate on more productive development efforts for the community such as work related to professional workstations and compatible Lisp support. The 2060 had a processing capacity of 2-3 times that of the dual KI-TENEX system, badly needed for our community, and it was more compact, reliable, and maintainable. Pending the arrival of more cost-effective and generally-available Lisp workstations, this would allow us to continue support for the SUMEX-AIM community at large and to provide facilities for new AI efforts.

In late 1982, we implemented the upgrade. The purchase price of the DECsystem 2060 reflected a substantial price reduction based on an external research grant from Digital Equipment Corporation to the Heuristic Programming Project in exchange for access by DEC to the AI software systems and knowledge-based systems expertise developed by the HPP. The remainder of the system was funded jointly by NIH and DARPA. The system configuration is shown in Figure 1. Of course, the transfer of service required a substantial investment of hardware engineering effort as all of the local line and network connections had to be changed over. This was all effected invisibly to the user community by running the old KI-TENEX and the new 2060 systems in parallel for more than a month.

Using DARPA funding, we also made some upgrades to the shared VAX 11/780 which was initially purchased by ARPA for HPP research as well as work in network graphics and VLSI design. The configuration of this machine is shown in Figure 3. In 1983, we augmented the machine by adding 2 Mbytes of memory and expanding the file system with a DEC RP07 disk drive (512 Mbytes). Approximately 60% of the machine is allocated for HPP and SUMEX use.

The overall facility model then became the central shared 2060, 2020, and VAX 11/780 systems surrounded with growing numbers of workstations and intercoupled by a local area network.

Additional Workstations

After the purchase of the 5 experimental Dolphin workstations, much work went into their development by Xerox, based on feedback and interactions with groups such as ours using them for AI applications. Performance of the Dolphins improved substantially based largely on improved microcoding of frequently used primitives and facilities. The initial optimizations of the Dolphin microcode were based on work at Xerox observing their own programs running. When the Dolphin was exposed to other AI systems such as ours, it became clear that additional improvements were necessary and were implemented, including enhanced performance for CONS operations, function calls, disk management, garbage collection, and other areas. Improvements in individual areas of performance ranged from factors of 2 to 10.

By 1983, other contenders were entering the Lisp workstation market in addition to Xerox. Because work in the HPP and the SUMEX-AIM community draws heavily on both Interlisp and the derivatives of MIT's MacLisp, we broadened our workstation experiments into both areas.

With NIH funding in 1983, we purchased 6 Xerox 1108 workstations (Dandelions) and in 1984, 3 Xerox 1109's (DandeTigers). With DARPA funding we purchased 2 Xerox 1108's and 1 1132 (high-performance Dorado) in 1984. In early 1985, the ONCOCIN group received a grant from Xerox of 13 1108's and additional printing and file server equipment. These machines represent the second generation of Xerox Lisp workstations and include significantly higher performance and functionality.

With DARPA funding in 1983 we bought a Symbolics LM-2 running the ZetaLisp system. In 1984, we added 3 Symbolics 3600's and a 3670 and in early 1985, another 3670 -- all with DARPA funding. We are also planning the purchase of additional workstations in the near term with DARPA funding.

Local Area Network Server Hardware

Since the late 1970's, we have been developing a local, high-speed Ethernet environment to provide a flexible basis for planned facility developments and the interconnection of a heterogeneous hardware environment. Our development of Ethernet facilities has been guided by the goals of providing the most effective range of services for SUMEX community needs while remaining compatible with and able to contribute to and draw upon network developments by other groups, dating back to the early 3 Mbit/sec Ethernet given to Stanford and several other universities by Xerox. We now support both 3 and 10 Mbit/sec Ethernets (see Figure 5) running numerous protocols and extended geographically throughout the SUMEX-AIM and related Stanford research groups. This network is the "glue" that holds the rest of the computing environment together and consists of numerous servers such as gateways and servers for terminal access, file storage and retrieval, and laser printing.

In the early phases, a substantial amount of special hardware was developed by our group for network interfaces including a high-performance direct memory access interface for the dual KI-TENEX system and a serial phase decoded UNIBUS interface that are used on our DEC 2020, VAX's, and early PDP-11 gateways and TIP's. The KI Ethernet interface served well for a period until we upgraded the system to a 2060, at which time we installed the 2060 mass bus EtherNet interface designed and built by the Stanford Computer Science Department. Our KI-10 interface is still seeing service in connecting another KI-10 system (Institute for Mathematical Studies in the Social Sciences) to the net.

Hardware for Gateways and TIP's

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. These first used PDP-11/05 hardware and then Motorola MC-68000 systems as they became available.

Similarly, we designed gateway between Apple equipment such as the new Macintosh terminal, that may play a role in our future virtual graphics work, and EtherNet using a MC-68000 gateway and a locally-designed Apple Bus to Multibus EtherNet interface. This system incorporates an 8530 Zilog chip to communicate with the Apple Net and software to manage the protocol packaging.

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. It can handle up to 32 lines. Both of these systems are now widely used throughout the Stanford network.

File Server Hardware

The development of an EtherNet file server was an integral part of our council-approved equipment plan with further expansions approved for later years. With joint NIH and DARPA funding, we were able to take advantage of an exceptional offer by Digital Equipment Corporation, through their corporate external research sponsorship program to DARPA contractors (the HPP), to purchase two VAX 11/750 machines as the processor part of our file servers. In the initial file server configurations, we also bought Fujitsu Eagle 450 MByte disks and controllers (one each from Systems Industries and Emulex) with one 800/1600 BPI tape unit for long term archives, and one 300 Mbyte removable pack drive for cyclic backups.

Other Network Hardware

We have developed numerous local network connection systems that have taken advantage of existing cabling rather than invest in expensive trenching and recabling. For example, in The Heuristic Programing Project (HPP) move to 701 Welch road, a high-performance network link to other SUMEX and campus network facilities was essential. Several communication schemes for establishing a reliable and relatively fast link were considered, including microwave, infrared laser, direct ethernet (by trenching and placing a direct ethernet cable), telephone company T1 service and others. All of these would have involved high cost and so we developed a communication link using bare copper telephone pair already in place. The wire distance between the HPP Welch Road location and the SUMEX machine room in the Medical Center is approximately 2000 ft. Utilizing high capacity differential drivers and ultra high speed, high sensitivity receivers, a half-duplex transceiver was developed for plain copper twisted pair that achieved error-free transmission at 1.25 Mbits/sec in each direction, utilizing Manchester data encoding. This communication link has been in operation for well over a year now without any appreciable down time or noticeable error rate or data delays.

In addition to the normal continuous flow of maintenance problems, we have reconnected the very reliable line printer from the old KI-TENEX system to the 2060. This required substantial modification of the printer controller to adapt to the different 2060 bus signal standards. We have also installed lots of communications equipment, including dial-in and -out modems and laser printer connections.

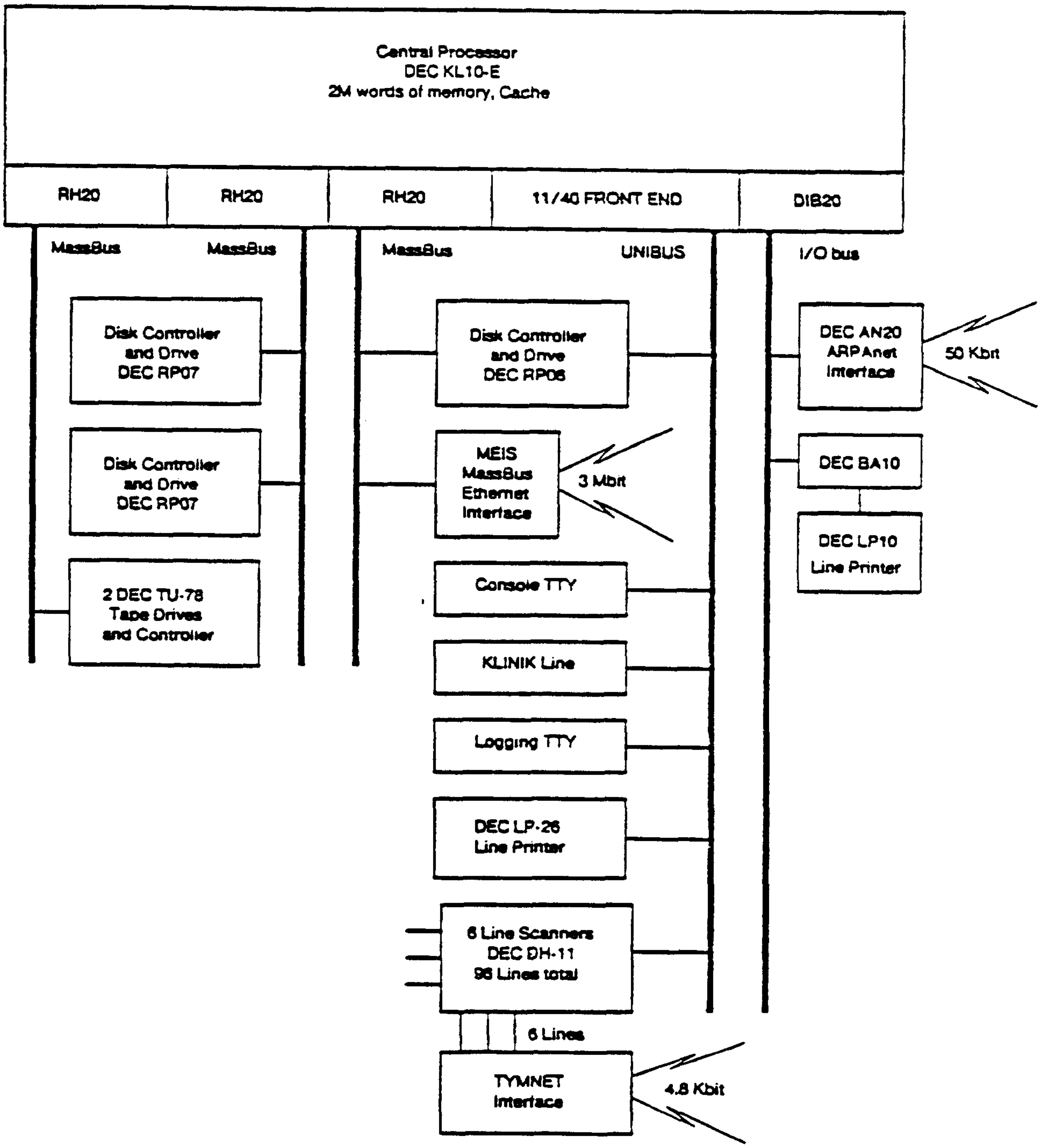

Figure 1: SUMEX-AIM DEC 2060 Configuration

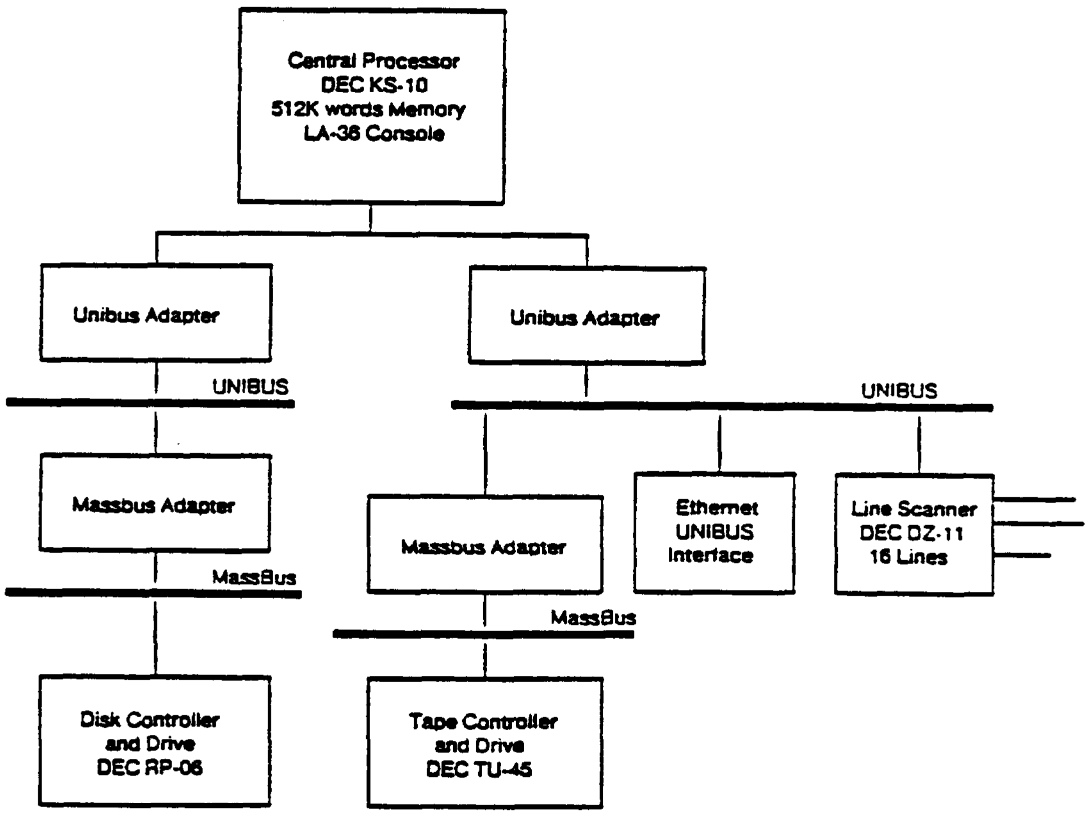

Figure 2: SUMEX-AIM DEC 2020 Configuration

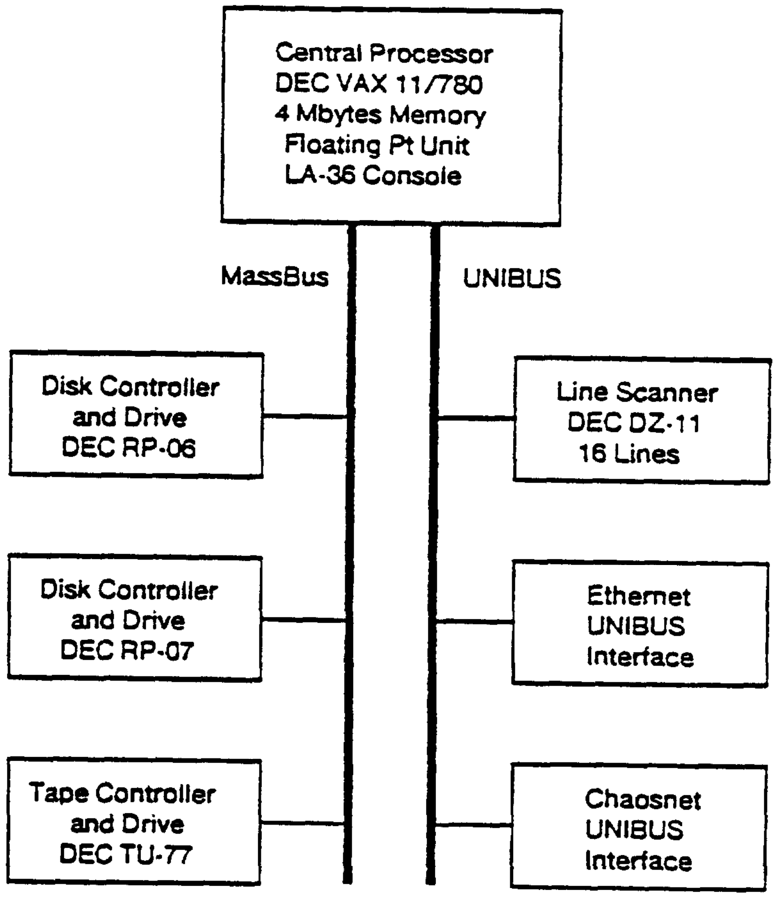

Figure 3: SUMEX-AIM Shared DEC VAX 11/780 Configuration

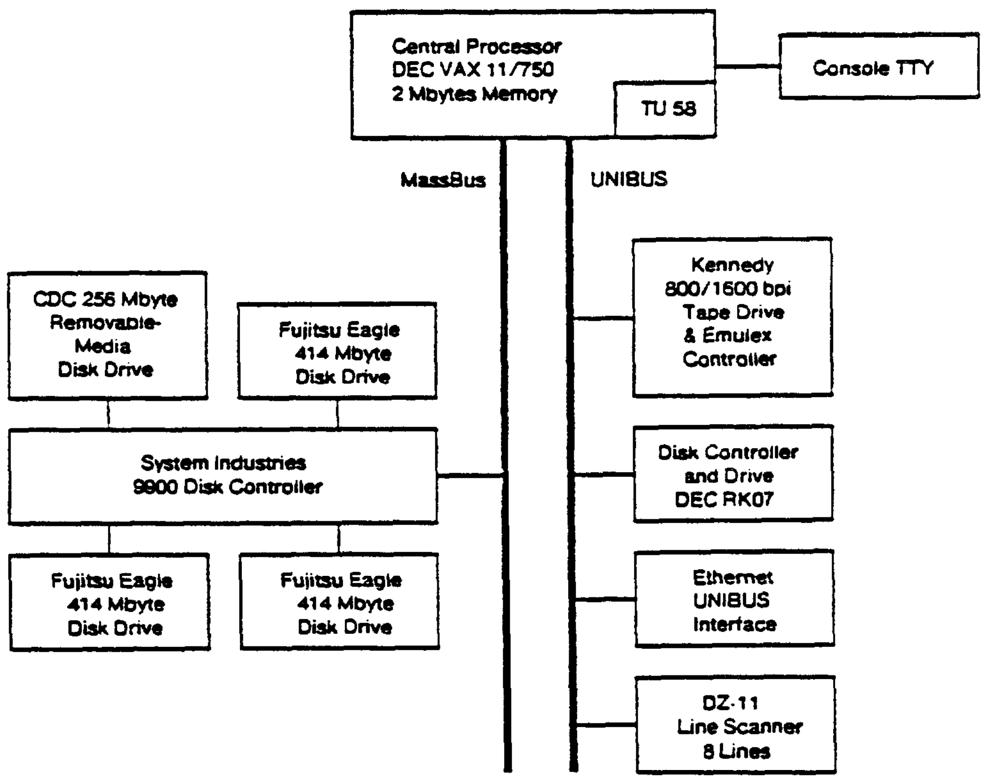

Figure 4: SUMEX-AIM File Server Configuration

III.A.3.3. Core System Development

Operating System Software

The various hardware elements of the SUMEX-AIM computing environment require the development and support of the operating systems that provide the interface between user software and the raw computing capacity. In addition to performance and relevance to AI research, much of our strategy for hardware selection has been based on being able to share development of the operating systems for our research among a large computer science community. This includes the mainframe systems (TOPS-20 and UNIX) and the workstation systems. Following are some highlights of recent system software developments.

TOPS-20 Development

The upgrade of the KI-TENEX system to the 2060 required a very large effort. Whereas the KI-TENEX system contained a great many local enhancements and adaptations, our goal was to run a TOPS-20 system that was broadly supported but which also tracked research developments outside of those motivated by vendor commercial interests. The most obvious choice for our immediate system peer community was the other 6 DEC 2060 sites at Stanford since we shared common internet problems and also had common goals in supporting research work rather than production computing. We also, of course, retained contact with the other ARPANET computer science systems. This course has constrained our own local developments by being part of a larger group of peers but the added problems of coordination have required fewer site-specific extensions and customizations at the operating system level.

Given this perspective, the following are specific areas of TOPS-20 system effort:

- In the conversion from TENEX, much planning and effort went into moving the file system, along with the pertinent user-specific directory information. In addition, we were able to preserve access to the vast magnetic tape library of archived and otherwise backed up files that had been created and saved since the inception of SUMEX. A TOPS-20 version of BSYS, a file archiving system, was imported from ISI as part of the effort to convert to the 2060. Numerous changes were made to make it compatible with the version of BSYS previously used at SUMEX. The LOOKUP program, used under TENEX, was converted to TOPS-20 use and made compatible with the new version of BSYS. We reviewed and updated appropriate documentation files in the HLP: and DOC: directories. And we identified and upgraded numerous system utility programs that utilized TENEX-dependent system calls.

- Using Tenex code previously developed at SUMEX as a base, we added new code to the TOPS-20 monitor to significantly enhance the user interface to the file system naming primitives. One addition was intercepting a ? typed by a user as part of a file name, then displaying for the user the valid file name alternatives matching the type-in up to that point, and finally returning to the original context, allowing the user to continue typing where he left off. Another addition was to generalize the logic involved in file name recognition in the case where more than one file matches what is typed in at the point where the request for recognition was given. The new logic looks ahead at the alternatives and fills out as much of the file name as possible, i.e. up to the point of ambiguity.

- Continued development of QANAL (formerly ANAL), a crash analysis program that has been under development since 1978. This program significantly eases the burden of analyzing the causes of system crashes due to both hardware and software problems. In addition, the accumulated outputs from QANAL allow for the detection of long term crash correlations to analyze infrequent problems.

- Track network protocol and service (e.g., file transfer and electronic mail) developments. We coordinated SUMEX's changes required to support the ARPANET-wide change from the old NCP protocols to the DOD IP/TCP protocols. This complex software required significant effort on our part because SUMEX-AIM has become a major communications crossroads and so exercises the network code very heavily. This has raised many problems of bugs and performance that we have worked to improve. We have played an active role in network discussion groups related to areas such as electronic mail, network designs, and protocols and had kept system tables for network host names and addresses, both local and over the ARPANET, up-to-date.

- Developed expanded file system support through multiple RP07 disk drive service. We were the first site to support more than one RP07 unit in a single structure.

- Implemented support for the old but superior LP10 printer from the KI-TENEX system. Even though DEC doesn't support this configuration, the LP10 has become our standard printer.

- Implemented subdirectory access to allow users full "owner" access to their subdirectories via the Access Control Job.

- Developed improved system allocation code, including the ability to withhold scheduler "windfall" from a given class or classes, with associated code in SKED% JSYS.

- Improved the efficiency of file backup and archive facilities by flagging directories with ARCHIVE and MIGRATE requests pending rather than searching through all directories serially.

- We have done substantial work on the TOPS-20 system Executive, the program that serves as the primary interface between users and the system. It provides commands to manipulate files, directories, and devices; control job and terminal parameter settings; observe job and system status; and execute public and private programs. The SUMEX EXEC is quite well developed at this stage but we have made several improvements. For example, we added a command line editor developed at the University of Texas and commands for the various laser printer spooling capabilities described later. There were also many more minor upgrades such as reading SYSTEM:LOGIN.CMD and SYSTEM:COMAND.CMD files on user login, account verification, enhancing various information commands, and improved directory and file system facilities to assist users in managing their files.

We have made numerous monitor bug and hardware problem repairs to provide for more reliable system operation and file integrity. Obvious bugs were removed long ago so those remaining are elusive and difficult to track down. We have also spent time keeping up-to-date with the latest monitor releases.

VAX 4.2 BSD UNIX Development

We run UNIX on our shared VAX 11/780 and on our 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. 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. The SUMEX VAX was the first site at Stanford to bring up the Berkeley 4.2 BSD distribution in October 1983. Since this was an early distribution, there were quite a number of bug fixes required; these were accomplished both through local effort and through monitoring the unix-wizards mailing list. After this kernel was running on the SUMEX machine, it was transported other sites and became the basis for the campus-wide UNIX 4.2 distribution.

To allow the UNIX network interface code to work in our Stanford subnet environment, we created a pseudo-network interface driver called 'sub0', that routed all output IP datagrams, based on their subnet numbers. This driver was done transparently, so that at system boot time, you could configure the machine for Stanford subnets, or for normal network routing. We also worked with other Stanford sites to install the Stanford PUP network drivers and servers back into 4.2 BSD (Berkeley does not support these).

Workstation System Development

Lisp workstations represent the major new direction for system development at SUMEX-AIM because these machines offer high performance Lisp engines, large address spaces required for sophisticated AI systems, flexible graphics interfaces for users, state-of-the-art program development and debugging tools, and a modularity that promises to be the vehicle for disseminating AI systems into user environments. We have accordingly invested a large part of our system effort in developing selected workstations and the related networking environments for effective use in the SUMEX-AIM community.

Xerox D-Machines

Much of the SUMEX-AIM community uses InterLisp and has moved naturally to the Xerox D-machines -- initially the Dolphin and then the Dandelion, Dandetiger, and Dorado. Much work has gone into hardware installation and networking support but we have also developed numerous software packages to help make the machines more effective for users and to ease our own problems in managing the distributed workstation environment.

In the transition to workstations as *computing* environments suitable for AI applications work, not just as programming environments, much system development remains to be done. One of the problems we have examined and plan to continue to exploring is that of building distributed expert systems. We are interested, for example, in separating the reasoning components and user interfaces and are designing a system with multiple processes which can run on a single or multiple workstations in order to independently develop, tune and evaluate the components. To facilitate this we have developed a prototype inter-process message passing interface which makes the topology of the system invisible to communicating processes, whether on one machine or several CPU's linked via the Ethernet.

Another of our interests is in exploring how to combine different software and/or hardware architectures in order to take advantage of the best features of each. One simple low level program that we built allows us to use Interlisp workstations to down load software into Mesa workstations in order to boot them using the Ethernet as an alternative to the hard or floppy disk drives. Along the same lines, we are exploring efficient ways to communicate high level descriptions of graphic data among differing media. We have developed a simple system which will take text formatting files and translate them into graphic window displays, defining active regions of the screen in the process. This facilitates the design of user interfaces using the familiar medium of text processing.

In our AI systems work, we have developed a low overhead object-oriented system which is designed to be flexible enough to model different object-oriented programming styles at the same time. It is also designed to facilitate a model of large knowledge bases which reside principally on file servers but whose components are loaded on demand. With this system, a minimal set of information about all the objects in a knowledge base is loaded upon opening. This information allows many simple inquires about the nature of objects and their relationships to be made without the main body of the object being resident. Only when non-trivial operations are performed are the contents of the object brought into core. This design is based on the belief that the size of knowledge bases will eventually grow to exceed the capacity of any given computer. However, most systems will generally only need a manageable subset of objects at runtime.

Other work we have done includes monitoring tools to examine static function calling hierarchy as well as view runtime executions graphically. We are also developing graphics interfaces to knowledge base construction and maintenance.

Some of the InterLisp software packages that have been written in the course of this work include:

ACFontCreate -- Reads a Xerox PARC font file in AC format into a lisp data structure

BaudRate -- Benchmarks baudrates by BINing through a file

DSys -- Monitors D machine usage on demand

GraphNet -- Derives topology of the PUP internet via net and gateway probes

HPColor -- Interlisp image stream implementation to drive H-P dgl graphics

Impress -- Interlisp image stream implementation to generate Impress print files

MakeStrike -- Writes out an Interlisp display font as a strike file

MLabel -- Generates mailing labels from a mailing list

RasterFontCreate -- Generates an Impress font of bitmap patches in arbitrary scale

ReadRSTFontFile -- Reads an Impress font file into a list data structure

RemoteTools -- Tools to manipulate a remote Interlisp using its systat process

RootPicture -- Reads a Press file bitmap into a lisp bitmap

RSTSample -- Creates an Impress sampler showing all characters of a font

SIL -- Reads and displays a SIL drawing file and optionally hardcopies it

SYSTAT -- a remote Eval server for Interlisp

Undither -- Compresses a previously dithered image into an AIS file

VDSDog -- Monitors array space usage to prevent crashing from lack thereof

WriteRSTFontFile -- generates an Impress font file from a special Lisp structure

ZDir -- TENEX-style directory lister for use with UNIX via Leaf server calls

DScribe -- A simple SCRIBE-to-display list parser/driver.

EtherBoot -- Provides microcode and program boot service for Xerox 8000's

GraphCalls -- Graphs the calling hierarchy of a lisp function and more

Hash -- Provide a machine independent hash file facility

EditBG -- A background/border texture editor.

FileLstW -- Menu-based interface to the file package.

MagnifyW -- A magnifying glass for bitmaps.

Message -- Multi-process/Multi-CPU message passing facility.

MultiW -- Links windows so that they move, surface, and close as a group

OZone -- An object-oriented programming system for Interlisp

Plotter -- Interlisp image stream to generate native-mode H-P plot files

Register -- Bundles menus into a coherent device for complex input

Region -- A utility to allow dissimilar activity in a single window.

Storage -- A utility to display Interlisp data type storage graphically.

Once a package has been developed and determined to be of general interest, we announce it over an electronic mail users list and make it available to other sites. In some cases, packages have such extensive utility that they are submitted as LispUsers packages for distribution by Xerox. This occurred in the case of Graphcalls, Hash, MultiW, and FileLstW, the latter submitted under the name Manager.

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.

Symbolics Lisp Machines

We have a growing community of Symbolics machines and users. Little development has gone into the tools for these systems yet because the small number of machines we have are concentrated in applications groups. We have actively supported the installation and maintenance of these systems, the installation of new software releases, and the integration of these systems with the rest of our networking environment. We were a beta test site for the Symbolics IP/TCP software.

Macintosh Workstations

In early 1984 Apple Computer released their new Macintosh and we were immediately interested in it as a possible low-cost display workstation to interface to our Lisp workstations and other hosts. In order to evaluate the Macintosh for this purpose, SUMEX received some early equipment and manuals through Stanford's participation in the Apple university consortium program. Like many groups trying to experiment with Macintosh software however, we found the Apple Lisa cross-development environment somewhat restrictive and hard to use and this was the only way to create Macintosh software at the time. So we built a UNIX-based cross-development environment on our VAX. It turns out, that this was the first C development environment available on the Macintosh when we released our software (via Arpanet FTP) in June of 1984. SUMacC (Stanford University Macintosh C) has been quite widely received, and is in use at well over a hundred sites throughout the US and in foreign countries. SUMACC integrated pieces of software from many groups, and was therefore something of a cooperative effort. We have openly distributed it to other users either through network FTP or a magnetic tape at distribution cost. Version 2.0 of the SUMACC system was released in November of 1984.

Among the many useful programs subsequently written with SUMACC were: (1) a Kermit program done at Harvard, (2) the Mac PSL (Portable Standard LISP) done at the University of Utah, and (3) an 'external file system' done by John Seamons of LucasFilm which allows the Macintosh to use an Ethernet host (such as UNIX) as a general network file server (see also page 37).

With the increased usage of Macintoshes in the SUMEX-AIM community, the need to be able to transfer files between them and TOPS-20 mainframes quickly arose. We therefore reimplemented the MACGet and MACPut file transfer utilities, previously developed for UNIX, for TOPS-20. These incorporated TOPS-20 style terminal handling and file system conventions. Both programs provide reliable (i.e., checksummed) transfer of either text or binary data, and are now gaining wide-spread use outside of SUMEX.

Virtual Workstation Graphics

Finally, we have done a number of experiments with the remote connection of bitmapped displays to hosts and workstations. Generally, the displays on Lisp machines are tethered through a high bandwidth cable to their processors. This limits the flexibility with which users can move from one Lisp machine to another (one must move physically to another machine) and loses the ability of researchers to work from home over telephone lines. A way of providing more flexible display to processor connection is to use a virtual graphics protocol, such as the V Kernel system developed by Lantz [18], that allows efficient communication of the contents to be displayed on a bitmapped screen. In an initial experiment, an Interlisp virtual graphics module was written to run on the DEC-2060 and drive the graphics engine of a Sun MicroSystems workstation over the Ethernet. This system allows applications running on the DEC-2060 to create views, and windows within those views on the remote workstation, and then using the Virtual Graphics Terminal Protocols, manipulate those views and windows. One can place text, draw objects such as points, lines, shaded rectangles, splines, and bitmaps in these screen areas. Local and remote editing of the graphics representation is also possible with a responsiveness close to that of a directly connected display.

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.

Remote Networks

We continue our connection to TYMNET as the primary means for access to SUMEX-AIM from research groups around the country and abroad. Substantial work was required to transfer TYMNET service from the KI-TENEX system to the 2060 because the new system does not support the same memory-sharing interface we had for the KI-10's. There has been no significant change in user service or network performance though. Very limited facilities for file transfer exist and no improvements appear to be forthcoming soon. Services continue to be purchased jointly with the Rutgers Computers in Biomedicine resource to maximize our volume usage price break. We continue to have serious difficulties getting needed service from TYMNET for debugging network problems and users away from major cities have problems with echo response times.

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. Terminal access restrictions are in force so that only users affiliated with DoD-supported contractors may use TELNET facilities. ARPANET is the primary link between SUMEX and other machine resources such as Rutgers-AIM and the large AI computer science community supported by DARPA. Our early Honeywell IMP has been upgraded to a BBN C/30 IMP in preparation for the transition to the IP/TCP protocols. We are also investigating the installation of 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.

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 and are now the standard system used in the campus-wide Stanford University network. We are developing gateways to interface other equipment as needed too (e.g., the Macintosh and Lisa). A diagram of our local area network system is shown in Figure 5 on 36 and the following summarizes our LAN-related development work.

MC-68000 Server Kernel -- Our early network gateways and TIP's were based on PDP-11 systems. But these soon became limiting in terms of speed, address space, and cost. With the introduction of the Motorola MC-68000 microprocessor and its integration into a compact, large-memory machine in the prototype SUN processor board developed in the Computer Systems Laboratory at Stanford, a much better vehicle was at hand. The net server software we developed for the PDP-11 included a kernel which handles hardware interfaces, core allocation, process scheduling, and low-level network protocol management. The 3 MBit/sec Ethernet PDP-11 based PUP kernel was translated and augmented for the MC-68000 CPU/SUN ethernet interface. This kernel then became the basis for the SUMEX gateway and TIP software which both have become the Stanford standard. As networking technology developed, the SUMEX kernel was extended to include 10 MBit/sec Ethernet drivers and to support 10 Mbit/sec PUP, XNS, and IP protocols. The main modification needed was the addition of a 10 MBit/sec Ethernet address resolution protocol module so that a 10 MBit/sec PUP host could discover its "soft" PUP address from a cooperating gateway on its local network.

Ethernet TIP -- Based on the new augmented MC-68000 kernel, the 3 Mbit/sec PDP-11 Ether TIP code was translated. This new TIP could handle increments of 8

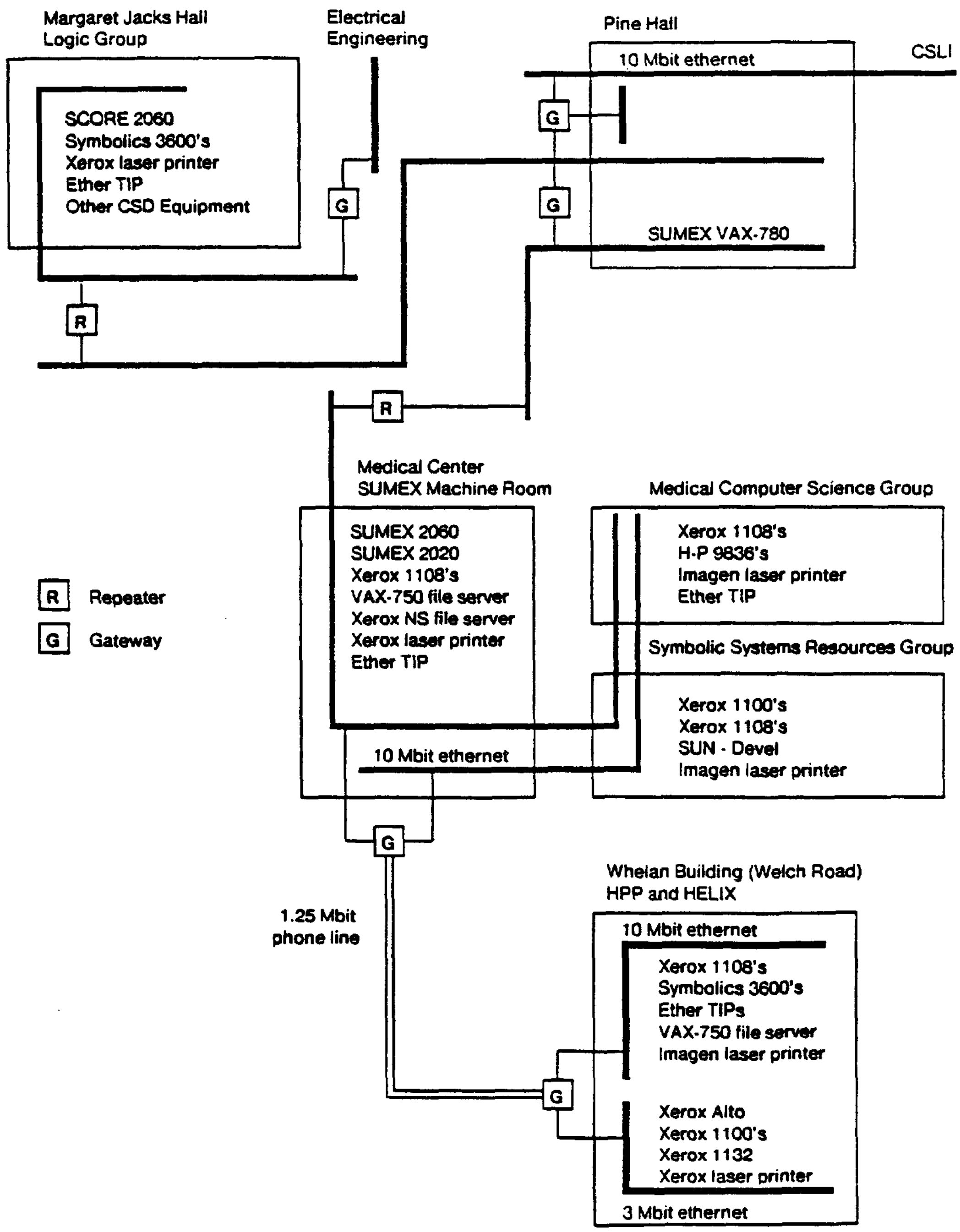

Figure 5: SUMEX-AIM EtherNet Configuration

lines up to 32 lines in a six slot backplane. With the advent of the newer 16 line DUART's developed in the Stanford Computer Science Department, 80 line TIP's have been built using this TIP code. This code is still running on several 3 Mbit/sec Ether TIP's at SUMEX. As 10 Mbit/sec networks were introduced, the TIP code was updated and adapted so that TIP's could run on either 3 MBit/sec or 10 MBit/sec Ethernets. There are now over 20 TIP's installed at Stanford using the SUMEX code and the number will increase substantially as the campus-wide local area network grows. The development of this software is essentially complete now with the recent addition of an improved user interface and facilities for inbound connections (such as for remote printers).

Ethernet Gateways -- Like the TIP systems, the PDP-11 gateway code was adapted to the MC-68000 hardware and extended to both 3 Mbit/sec and 10 Mbit/sec networks. Gateways can be configured to support up to four directly connected networks which may be either 3 MBit/sec or 10 MBit/sec. The gateway system was made "self-configuring" so that only one bootable gateway was needed. Network directory downloading and name/address lookup services were added. The routing algorithm was rewritten to minimize probe time for efficiency because of the continued growth of the number of subnetworks in the Stanford University network. The gateway now supports PUP and IP packet transport and XNS packet routing for both 10mb and 3mb networks is being completed. There are over twenty SUMEX gateways installed at Stanford and this number should double in the next year.

A special gateway configuration was required for the HPP move to Welch Road. Since the physical link was differentially driven 1.25 MBit/sec twisted pair cable, the network connections required two three-way gateways, one at either end, and special hardware to interface the serial lines with the ethernet interfaces. The required special hardware and software were built and the WR gateway has operated very effectively.

Apple Gateway Another special gateway, named SEAGATE, was developed to better integrate the Apple Macintosh into our Ethernet system. It links the Ethernet and Apple's AppleBus/AppleTalk network. This was completed and released in February 1985. Several internet sites, including some at Stanford, are currently constructing duplicate gateways. Also, several commercial firms are building a one board version of the gateway which should lower the cost to about $1000 per gateway. EFS, MAT, and AppleTalk Library are some sample Macintosh programs and UNIX daemons, that utilize SEAGATE. EFS is an external file system, written by John Seamons, and modified by us to work over AppleTalk. With EFS the Mac user sees his normal iconic view of the world. His UNIX directory appears as an icon and he can remotely execute and transfer files, simply by clicking on their icons. EFS is to the Mac as Leaf is to a LISP machine. The AppleTalk library is used by all of these programs to perform the ATP protocol (AppleTalk transaction protocol). This is the general protocol used to perform printing, file transfer, etc. with the Mac. The library allows a UNIX user-level process to perform this ATP protocol. Note that no kernel changes are required, since the ATP datagrams are imbedded in IP datagrams (UDP) by the SEAGATE. MAT is the Mac ATP Transfer program, a sample program that does file transfers with a UNIX host. It can also act as the framework for a Mac mail or print service.

Remote File Service -- In a distributed workstation environment, effective file access and transfer facilities between workstations and other hosts and servers are a must, especially to file servers like those we built around VAX 11/750 UNIX systems. Initial file service support used code written as a student project in the Stanford Computer Systems Laboratory. But as the number of workstations increased, service degraded and it became necessary to rewrite the PUP/BSP UNIX software package, and major portions of those programs dependent upon these protocols. This resulted in a 300% increase in throughput and stabilized the Lisp Machine to VAX 11/750 file service environment. At the same time we made major improvements to the UNIX Leaf service for XEROX D-machines. The earlier code, again a student systems project, had many bugs and inefficiencies and required a complete rewrite. In the new code, each Leaf connection was given a separate process to manage its Leaf resources, whereas previously, all users' Leaf requests were simply handled as a serial queue. This meant that every packet created a bottleneck for its successors. This work resulted in a much better Leaf service environment with considerable improvement in overall responsiveness and throughput.

Laser Printing Services

Since the first Xerox laser printers were developed in the mid-1970's, several companies have produced computer-driven systems, such as the Xerox Raven and the Imagen 8/300. These systems have become essential components of the work 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.

We collaborated to develop an Ethernet interface for Imagen printers starting about January of 1984. We arranged to upgrade our Imprint-10 controller in exchange for the UNIX software needed to drive it from the network and were the first site to receive this controller in beta test stage. The UNIX software we developed made it possible to connect the printer to the new 4.2 BSD line printer spooler package using IP/TCP protocols. This was completed about March of 1984. After the UNIX implementation was complete, we developed the corresponding TOPS20 software to interface to this new printer and later, integrated it into the TOPS20 Galaxy spooler package. Other sites on campus and in the internet, began using the new printer and our spooling software as well.

We similarly developed and enhance the spooling system for the Dover and Alto-Raven laser printers and added a header page for Raven output to separate listings. And in addition to the device support for the printers to interface to the various mainframe hosts machines in our network, we also developed packages to allow Xerox D-machines and Symbolics 3600 machines to print to the networked laser printers.

On the SUMEX-AIM mainframe hosts, SCRIBE is the predominant document compilation system, but in the initial stages, it was essentially only used with the Xerox Dover printer or a daisywheel typewriter. In the succeeding years we integrated the Imagen Imprint-10 driver from Unilogic, brought up the Xerox Alto-Raven, and installed support for the new group of Imagen printers (the 8/300's), which are based on a Canon copier and are now the workhorse printing resources of the local community. We made numerous improvements in the printing fonts available to users, including a rework of Knuth's Computer Modern Roman fonts for a more contemporary look on the Imprint-10, creating a sans serif font family based on Computer Modern Roman, generating Helvetica and Times Roman font families from the Xerox sources used to generate the Dover fonts, and creating and improving many document types in use by the community.

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, DEC-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 on operating systems and networking. A further gratifying example is the TTYFTP program, originally written at SUMEX as a system for file transfers usable over any circuit that appears as a terminal line to the operating system (hardline, dial-up, TYMNET, etc.) and incorporating appropriate control protocols and error checking. The design was derived from the DIALNET protocols developed at the Stanford AI Laboratory with extensions to allow both user and server modules to run as user processes without operating system changes. TTYFTP formed the basis for the KERMIT program that is now distributed by Columbia University and which is in very wide use for communications between personal computers and to mainframe hosts.

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 81).

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 upgrade of the KI-TENEX system to the 2060 required major planning and care to ensure continuous resource operation during the phase-over. Similarly, the relocation of our VAX 11/780 to Pine Hall and the outfitting of the KSL machine room at the Welch Road laboratory required much effort.

We use students for much of our operations and related systems programming work. Over the past 4 years, we have hired and trained a total of 15 undergraduate operations assistants.

We also 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.

III.A.3.4. Core AI Research

We have maintained a strong core AI research effort in the SUMEX-AIM resource aimed at developing information resources, basic AI research, and tools of general interest to the SUMEX-AIM community. It should be noted that the SUMEX resource grant from NIH provides much of the computing environment for this core AI work[1] but NIH supports only a small part of the manpower and other support for core AI. For example, NIH has provided partial funding for work on the AI Handbook, the AGE project, and part of the core ONCOCIN development for the dissemination of consultative AI systems. Substantial additional support for the personnel costs of our core AI research (roughly comparable to the NIH investment in computing resources) comes from DARPA, ONR, NSF, NASA, and several industrial basic research contracts to the Knowledge Systems Laboratory or KSL[2] (see the summary of core research funding on page 47).

Our core AI research work has long been the mainstay on which our extensive list of applications projects are based. This work has been focused on medical and biological problems for over a decade with considerable success, particularly in the area of expert systems which represent one important class of applications of AI to complex problems -- in medicine, science, engineering, and elsewhere. Numerous high-performance, expert systems have resulted from our work on expert systems in such diverse fields as analytical chemistry, medical diagnosis, cancer chemotherapy management, VLSI design, machine fault diagnosis, and molecular biology. Other projects have developed generalized software tools for representing and utilizing knowledge (e.g., EMYCIN [4, 34], UNITS [33], AGE [25], MRS [9], GLISP [27]) as well as comprehensive publications such as the three-volume *Handbook of Artificial Intelligence* [1] and books summarizing lessons learned in the DENDRAL [21] and MYCIN [4, 32] research projects.

But the current ideas fall short in many ways, necessitating extensive further basic research efforts. Our core research goals 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.

The following summary reports progress on the basic or core research activities within the KSL. As indicated earlier, the development of the ONCOCIN system (under Professor Shortliffe) is an important part of our core research proposal for the renewal period. Progress on that work is reported separately in Section IV.A.3 on page 102, however, because its efforts have been supported as a collaborative and resource-related research project up until now. Together, this work explores a broad range of basic research ideas in many application settings, all of which contributes in the long term to improved knowledge based systems in biomedicine.

Recent Highlights of Research Progress

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 Hodgkins disease. Thus, within the KSL

---

[1] DARPA funds have also helped substantially in upgrading the KI-TENEX system to the 2060 and in the purchase of community Lisp workstations

[2] See Appendix A on page 203 for an overview of the KSL organization.

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.

In addition to the technical reports listed later, the following books and survey articles were published just during this year -- 11 books total have been published in the past 4 years as indicated in Appendix A. These are of central interest to AI researchers and of direct relevance to the mission of the SUMEX-AIM resource.

BOOKS:

1. Buchanan, B.G. and Shortliffe, E.H., eds. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Reading, MA: Addison-Wesley Publishing Company, 1984.
2. Clancey, W.J. and Shortliffe, E.H., eds. *Readings in Medical Artificial Intelligence: The First Decade.* Reading, MA: Addison-Wesley Publishing Company, 1984.
3. Cohen, Paul R. *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach.* London and Marshfield, MA: Pitman Advanced Publishing Program, 1985.

SURVEY ARTICLES: HPP 84-15, 84-20, 84-23, 84-28, and 84-32.

In addition, work is progressing on a textbook for students beginning to study medical computing and artificial intelligence[1]. This multi-authored volume should be completed in draft form by the end of 1985 and a 1986 publication date is contemplated. Writing this new book will be facilitated by the SUMEX resource, much as the *Handbook of AI* was in the past. A multi-authored text of this type, particularly one for which the authors are spread at numerous different universities around the country, would be a nightmare to compile if it were not for the SUMEX resource. Many of the contributors to the book have been assigned SUMEX accounts for purposes of manuscript preparation. On-line manuscript work through the shared facility, coupled with messaging capabilities, will greatly enhance the efficiency and accuracy of the developing chapters and the editing process.

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

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

   A working version of NEOMYCIN has been implemented which demonstrates the effectiveness of representing strategy knowledge explicitly. A detailed study of rule-based systems was published in book form. Specific representational issues in logic-based systems were addressed in the

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

context of MRS. We designed a method for representing temporal knowledge in ONCOCIN. Finally, Cooper's Ph.D. thesis on representing and using causal and probabilistic knowledge was published in this year.

[See KSL technical memos KSL-84-9, KSL-84-10, KSL-84-18, KSL 84-31, KSL-84-41, KSL-85-5.]

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

   Much of our research in the past year has involved investigations with the Blackboard architecture begun in previous years. We have implemented our design in a working system called BB1.

   [See KSL technical memos KSL-84-11, KSL-84-12, KSL-84-14, KSL 84-16, KSL 84-36.]

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

   Three Ph.D. theses (Fu, Greiner, and Dietterich) in the area of knowledge acquisition were completed in this year. Fu's work develops methods for learning by induction, where the target rules may have some associated degrees of uncertainty and may contain names of intermediate concepts. This work was demonstrated in the context of diagnosing causes of jaundice. Greiner's work examines learning by analogy. Dietterich's work elucidates methods needed in learning programs to deal with state variables and with problems of using a partially learned theory to interpret new data that will be used to learn new elements of the theory. In addition, we implemented the first parts of a program that can learn by watching an expert. And we implemented a prototype system that learns control heuristics from an expert using a problem solving program written in BB1.

   [Preliminary results have been published in KSL-84-10, KSL-84-18, KSL-84-24, KSL-84-38, KSL-84-45, KSL 84-46, KSL-85-2, KSL-85-4.]

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

   We completed the design of a system using Dempster's rule of propagating uncertainty, and we examined several other issues regarding the use of probabilistic information in expert systems. Dr. Jean Gordon, a mathematician and Stanford medical student, collaborated with Dr. Shortliffe on work that examines inexact inference using the Dempster-Shafer theory of evidence, demonstrating its relevance to a familiar expert system domain, namely the bacterial organism identification problem that lies at the heart of the MYCIN system, and presenting a new adaptation of the D-S approach with both computational efficiency and permitting the management of evidential reasoning within an abstraction hierarchy.

   We examined the use of counter-factual conditionals in logic-based systems and completed an analysis of how procedural hints can be used by a problem solver.

[See KSL technical memos KSL-84-11, KSL-84-17, KSL-84-21, KSL-84-30, KSL-84-31, KSL-84-35, KSL 84-41, KSL-84-42, KSL-84-42, KSL-84-43.]

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

   We have continued the development of new software tools for expert system construction and the distribution of packages that are reliable enough and documented so that other laboratories can use them. These include the old rule-based EMYCIN system, MRS, and AGE. Progress has been made in making the BB1 instantiation of the blackboard architecture domain-independent. We have begun constructing and editing subsystems and have completed a first implementation of an explanation subsystem.

   [See KSL technical memos KSL-84-16, KSL-84-39.]

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

   A program for inferring a model of users was designed and implemented in the context of a tutoring system that aids in teaching algebra. A second user-modelling program was implemented in the context of NEOMYCIN to help understand how an expert solves problems. A survey of explanation capabilities in medical consultation programs was published.

   A new project on knowledge-based explanations in a decision analysis environment is getting underway as the thesis research of Dr. Glenn Rennels. This work is actually a synthesis of artificial intelligence, decision analysis and statistics. The work concerns medical management, not diagnosis; diagnostic decisions identify underlying mechanisms of the illness, and group the patient's problems under a diagnostic label, whereas management decisions plan actions that will prevent undesirable outcomes and restore health. The intelligent behavior we want to emulate is (a) the identification of studies relevant to a given clinical case, and (b) interpretation of those studies for decision-making assistance.

   [See KSL technical memos KSL-84-12, KSL 84-27, KSL-84-29.]

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

   A major paper on skeletal planning was published in this year. And we published in the biochemistry literature some results of applying skeletal planning to experiment design in genetic engineering.

   [See KSL technical memos KSL-84-33, KSL-85-6.]

8. *Diagnosis*: How can we build a diagnostic system that reflects any of several diagnostic strategies? How can we use knowledge at different levels of abstraction in the diagnostic process?

   Research on using causal models in a medical decision support system (NESTOR) was published in this year. Using the domain of hypercalcemic disorders, NESTOR attempts to use knowledge-based methods within a formal probability theory framework. The system is able to score hypotheses with causal knowledge guiding the application of sparse probabilistic knowledge; search for the most likely hypothesis without exploring the entire hypothesis space; and critique and compare hypotheses which are generated by the system, volunteered by the user, or both.

A second medical diagnosis program that uses causal models of renal physiology (AI/MM) was also published. In this system, analysis and explanation of physiological function is based on two kinds of causal relations: empirical "Type-1" relations based on definitions or on repeated observation and mathematical "Type-2" relations that have a basis in physical law. Inference rules are proposed for making valid qualitative causal arguments with both kinds of causal basis.

A working implementation of the PATHFINDER system was evaluated and its diagnostic strategies were analyzed. A taxonomy of diagnostic methods was completed and integrated into the NEOMYCIN system.

[See KSL technical reports: KSL-84-13, KSL-84-19, KSL-84-48, KSL-85-5.]

Relevant Core Research Publications

HPP 84-9 David H. Hickam, Edward H. Shortliffe, Miriam B. Bischoff, A. Carlisle Scott, and Charlotte D. Jacobs; *Evaluations of the ONCOCIN System: A Computer-Based Treatment Consultant for Clinical Oncology, (1) The Quality of Computer-Generated Advice and (2) Improvements in the Quality of Data Management,* May 1984.

HPP 84-10 Thomas G. Dietterich; *Learning About Systems That Contain State Variables,* June 1984. In *Proceedings of AAAI-84,* August 1984.

HPP 84-11 M. Genesereth, and D.E. Smith; *Procedural Hints in the Control of Reasoning,* May 1984.

HPP 84-12 Derek H. Sleeman; *UMFE: A User Modelling Front End Subsystem,* April 1984.

HPP 84-13 Eric J. Horvitz, David E. Heckerman, Bharat N. Nathwani, and Lawrence M. Fagan; *Diagnostic Strategies in the Hypothesis-Directed PATHFINDER System,* June 1984, submitted to the *First Conference on Artificial Intelligence Applications, Denver, CO.,* December 5-7, 1984.

HPP 84-14 Vineet Singh, and M. Genesereth; *A Variable Supply Model for Distributing Deductions,* May 1984.

HPP 84-15 Bruce G. Buchanan; *Expert Systems,* July 1984, *Journal of Automated Reasoning, Vol. 1, No. 1, Fall,* 1984.

HPP 84-16 *STAN-CS-84-1034* Barbara Hayes-Roth; *BB-1: An Architecture for Blackboard Systems That Control, Explain, and Learn About Their Own Behavior,* December 1984.

HPP 84-17 M.L. Ginsberg; *Analyzing Incomplete Information,* 1984.

HPP 84-18 William J. Clancey; *Knowledge Acquisition for Classification Expert Systems,* July 1984, *Proceedings of ACM-84,* 1984.

HPP 84-19 E.H. Shortliffe; *Coming to Terms With the Computer,* to appear in S.R. Reiser, and M. Anbar (eds.), *The Machine at the Bedside: Strategies for Using Technology in Patient Care, Cambridge University Press,* 1984.

HPP 84-20 E.H. Shortliffe; *Artificial Intelligence and the Future of Medical Computing,* in *Proceedings of a Symposium on Computers in Medicine, annual meeting of the California Medical Association, Anaheim, CA.,* February 1984.

HPP 84-21 E.H. Shortliffe; *Reasoning Methods in Medical Consultation Systems: Artificial Intelligence Approaches (Tutorial),* in *Computer Programs in Biomedicine* January 1984.

HPP 84-22 *ONCOCIN Project: Studies to Evaluate the ONCOCIN System; 6 Abstracts,* February 1984.

HPP 84-23 Edward H. Shortliffe; *Feature Interview: On the MYCIN Expert System,* in *Computer Compacts, 1:283-289,* December 1983/January 1984.

HPP 84-24 B.G. Buchanan, and E.H. Shortliffe; *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project,* published with *Addison-Wesley, Reading, MA.,* 1984.

HPP 84-25 W.J. Clancey, and E.H. Shortliffe; *Readings in Medical Artificial Intelligence: The First Decade,* published with *Addison-Wesley, Reading, MA.,* 1984.

HPP 84-27 Edward H. Shortliffe; *Explanation Capabilities for Medical Consultation Systems (Tutorial),* in D. Lindberg, and M. Collen (eds.), *Proceedings of AAMSI Congress 84, pp. 193-197, San Francisco,* May 21-23, 1984.

HPP 84-28 E.H. Shortliffe, and L.M. Fagan; *Artificial Intelligence: The Expert Systems Approach to Medical Consultation,* in *Proceedings of the 6th Annual International Symposium on Computers in Critical Care and Pulmonary Medicine, Heidelberg, Germany,* June 4-7, 1984.

HPP 84-29 David C. Wilkins, Bruce G. Buchanan, and William J. Clancey: *Inferring an Expert's Reasoning by Watching, Proceedings of the 1984 Conference on Intelligent Systems and Machines,* 1984.

HPP 84-30 M.L. Ginsberg: *Non-Monotonic Reasoning Using Dempster's Rule,* June 1984.

HPP 84-31 M.L. Ginsberg: *Implementing Probabilistic Reasoning,* June 1984.

HPP 84-32 Bruce G. Buchanan: *Artificial Intelligence: Toward Machines That Think,* July 1984, in *Yearbook of Science and the Future, pp. 96-112, Encyclopedia Britannica, Inc., Chicago,* 1985.

HPP 84-33 Rene Bach, Yumi Iwasaki, and Peter Friedland; *Intelligent Computational Assistance for Experiment Design,* in *Nuclear Acids Research,* January 1984.

MCS Thesis Kunz, John C.; *Use of Artificial Intelligence and Simple Mathematics to Analyze a Physiological Model,* Doctoral dissertation, Medical Information Sciences, June 1984.

HPP 84-35 Jean Gordon, and Edward Shortliffe; *A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space,* September 1984 and in *Artificial Intelligence,* 26(3), July 1985.

HPP 84-36 Michael R. Genesereth, Matt Ginsberg, and Jeff S. Rosenschein; *Cooperation Without Communication,* September 1984.

HPP 84-38	Li-Min Fu, and Bruce G. Buchanan; *Enhancing Performance of Expert Systems by Automated Discovery of Meta-Rules,* September 6, 1984.
HPP 84-39	Paul S. Rosenbloom, John E. Laird, John McDermott, Allen Newell, and Edmund Orciuch; *R1-Soar: An Experiment in Knowledge-Intensive Programming in a Problem-Solving Architecture,* to appear in the *Proceedings of the IEEE Workshop on Principles of Knowledge-Based Systems,* October 1984.
HPP 84-41	*STAN-CS-84-1032* Michael R. Genesereth, Matthew L. Ginsberg, and Jeffrey S. Rosenschein; *Solving the Prisoner's Dilemma,* November 1984.
HPP 84-42	Matthew L. Ginsberg; *Does Probability Have a Place in Non-Monotonic Reasoning?* submitted to the *IJCAI-85,* November 1984.
HPP 84-43	*STAN-CS-84-1029* Matthew L. Ginsberg; *Counterfactuals,* submitted to the *IJCAI-85,* December 1984.
HPP 84-45	Devika Subramanian, and Michael R. Genesereth; *Experiment Generation with Version Spaces,* December 1984.
HPP 84-46	Thomas G. Dietterich; *Constraint Propagation Techniques for Theory-Driven Data Interpretation,* PhD Thesis, to be published as a book by *Kluwer,* December 1984.
HPP 84-48	*STAN-CS-84-1031* Gregory F. Cooper; *NESTOR: A Computer-Based Medical Diagnostic Aid That Integrates Causal and Probabilistic Knowledge,* PhD Thesis, December 20, 1984.
KSL 85-2	*STAN-CS-85-1036* Barbara Hayes-Roth, and Michael Hewett; *Learning Control Heuristics in BB1,* submitted to the *IJCAI-85,* January 1985.
KSL 85-4	(Needs Authors Permission) Li-Min Fu, and Bruce G. Buchanan; *Learning Intermediate Knowledge in Constructing a Hierarchical Knowledge Base,* submitted to the *IJCAI Conference Proceedings for 1985,* January 1985.
KSL 85-5	(Needs Authors Permission) William J. Clancey; *Heuristic Classification,* March 1985.
KSL 85-6	Peter E. Friedland, and Yumi Iwasaki; *The Concept and Implementation of Skeletal Plans,* published in the *Journal of Automated Reasoning,* 1985.
KSL 85-7	Rene Bach, Yumi Iwasaki, and Peter Friedland; *Intelligent Computational Assistance for Experiment Design,* published in *Nucleic Acids Research,* 1985.
KSL 85-8	(Needs Authors Permission) M.G. Kahn, J. Ferguson, E.H. Shortliffe, and L. Fagan; *An Approach for Structuring Temporal Information in the ONCOCIN System,* March 1985.

Summary of Core Research Funding Support

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

Interactions with the SUMEX-AIM Resource

Our interactions with the SUMEX-AIM resource involve the facilities -- both hardware and software -- and the staff -- both technical and administrative. Taken together as a whole resource, they constitute an essential part of the research structure for the KSL. Many of the grants and contracts from other agencies have been awarded partly because of the cost-effectiveness of AI research in the KSL due to the fact that much of our computing needs could be more than adequately met by the SUMEX-AIM resource. In this way the complementary funding of this work by the NIH and other agencies provides a high leverage for incremental investment in AI research at the SUMEX-AIM resource.

We rely on the central SUMEX facility as a focal point for all the research within the KSL, not only for much of our computing, but for communications and links to our many collaborators as well. As a common communications medium alone, it has significantly enhanced the nature of our work and the reach of our collaborations. The existence of the central time-shared facility has allowed us to explore new ideas at very small incremental cost.

As SUMEX and the KSL acquire a diversity of hardware, including LISP workstations and smaller personal computers, we rely more and more heavily on the SUMEX staff for integration of these new resources into the local network system. The staff has been extremely helpful and effective in dealing with the myriad of complex technical issues and leading us competently into this world of decentralized, diversified computing. At the same time, the staff has provided a stable, efficient central time-shared machine running software that has been developed at many sites over many years. Without the dedication of the SUMEX staff, the KSL would not be at the forefront of AI research.

III.A.3.5. Training Activities

The SUMEX resource exists to facilitate biomedical artificial intelligence applications from program development through testing in the target research communities. 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 system and use it effectively. 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.

On another front, 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.

And finally, the training of graduate students is an essential part of the research and educational activities of the KSL. Currently 41 students are working with our projects centered in Computer Science and another 20 students are working with the Medical Computer Science program in Medicine. Of the 41 working in Computer Science, 25 are working toward Ph.D. degrees, and 16 are working toward M.S. degrees. A number of students are pursuing interdisciplinary programs and come from the Departments of Engineering, Mathematics, Education, and Medicine.

Based on the SUMEX-AIM community environment, we have initiated two unique and 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, who are knowledgeable about knowledge-based system techniques.

The *Medical Information Sciences (MIS)* program is 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. The MIS training program is based in School of Medicine, directed by Dr. Shortliffe, co-directed by Dr. Fagan, and overseen by a group of nine University faculty that includes several faculty from the Knowledge Systems Laboratory (Profs. Shortliffe, Feigenbaum, Buchanan, and Genesereth). It was Stanford's active ongoing 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 focusses 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 a second class of five entered last summer. A third group of seven students has just been selected to begin during 1985. The proposed steady state size for the program (which should be reached in 1986) is 20-22 trainees. Applicants to the program in our first two years have come from a number of backgrounds (including seven MD's and five medical 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.

With the exception of this latter group, 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 in this application.

Early evidence suggests that the program already has an excellent reputation due to:

- high quality students, many of whom are beginning to publish their work in conference proceedings and refereed journals;
- 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 (almost $1M in hardware gifts from Hewlett Packard, Xerox, and Texas Instruments; over $200K in cash donations from corporations and foundations; and an NIH post-doctoral training grant from the National Library of Medicine).

The *Master of Science in Computer Science: Artificial Intelligence (MS:AI)* program is a terminal professional degree offered for students who wish to develop a competence in the design of substantial knowledge-based AI applications but who do not intend to obtain a Ph.D. degree. The MS:AI program is administered by the Committee for Applied Artificial Intelligence, composed of faculty and research staff of the Computer Science Department. Normally, students spend two years in the program with their time divided equally between course work and research. In the first year, the emphasis is on acquiring fundamental concepts and tools through course work and and project involvement. During the second year, students implement and document a substantial AI application project.

III.A.3.6. Resource Operations and Usage

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

1. Overall resource loading data (page 53).
2. Relative system loading by community (page 54).
3. Individual project and community usage (page 57).
4. Network usage data (page 64).
5. System reliability data (page 64).

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 severe changes in most of the graphs. This is due to many reasons briefly mentioned here:

1. Even though the TENEX operating system used on the KI-10 was a forerunner of the current Tops20 operating system, the Tops20 system is still different from TENEX is many ways. Tops20 uses a radically different job scheduling mechanism, different methods for computing monitor statistics, different I/O routines, etc. In general, it can not be assumed that statistics measured on the TENEX system correlate one to one with similar statistics under Tops20.
2. The KL-10 processor on the 2060 is a faster processor than the KI-10 processor used previously. Hence, a job running on the KL-10 will use less CPU time than the same job running on the KI-10. This aspect is further complicated by the fact that the SUMEX KI-10 system was a dual processor system.
3. The SUMEX-AIM Community was changing during the time of the transfer to the 2060. The usage of the GENET community on SUMEX had just been phased out. This part of the community accounted for much of the CPU time used by the AIM community. Since the purchase of the 2060 was partially funded by the Heuristic Programming Project (HPP), an additional number of HPP Core Research Projects started using the 2060, increasing the Stanford communities usage of the machine. And finally, the move to the 2060 occurred during a pivotal time in the community when more and more projects were either moving to their own local timesharing machines, or onto specialized Lisp workstations. It also was the time for the closure of many long time SUMEX-AIM projects, like DENDRAL and PUFF/VM.

Any conclusions reached by comparing the data before and after February, 1983 should be done with caution. The data is included in this years annual report mostly for casual comparison.

Also, it should be noted that monthly statistics are not available for this past year because of problems with the accounting program at this writing. The appropriate average data quantity for the year is shown instead for each month so the graphs appear to be "flat" in the area corresponding to the current period.

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.

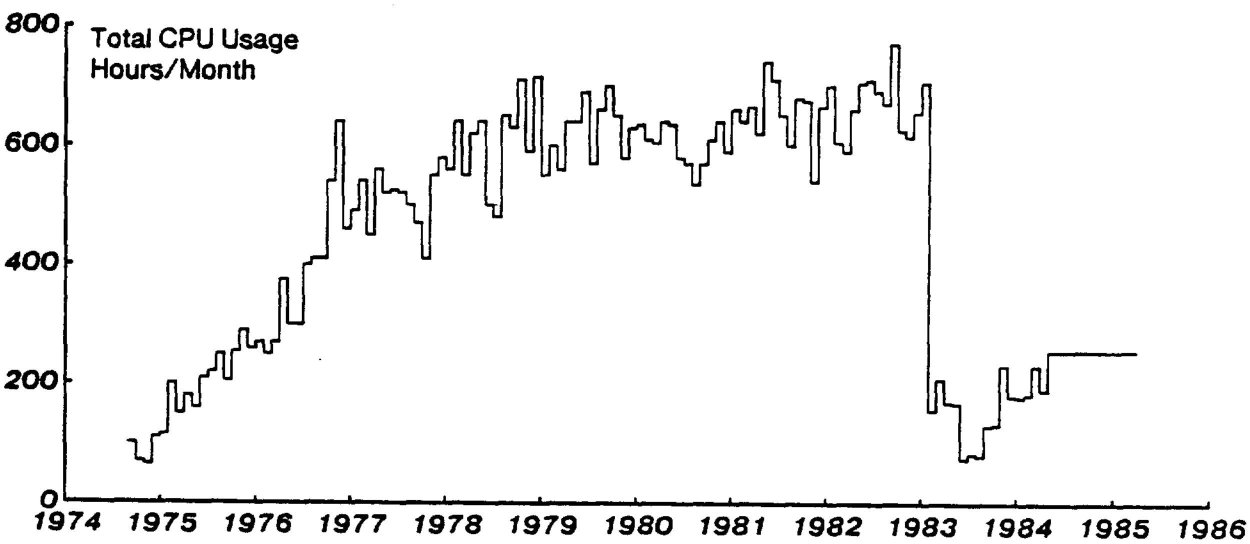

Figure 6: Total CPU Time Consumed by Month

Relative System Loading by Community

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

Stanford	40%
AIM	40%
Staff	20%

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

The monthly usage of CPU resources and terminal connect time for each of these three communities relative to their respective aliquots is shown in the plots in Figure 7 and Figure 8. As mentioned on page 52, these plots include both KI-10 and 2060 usage data.

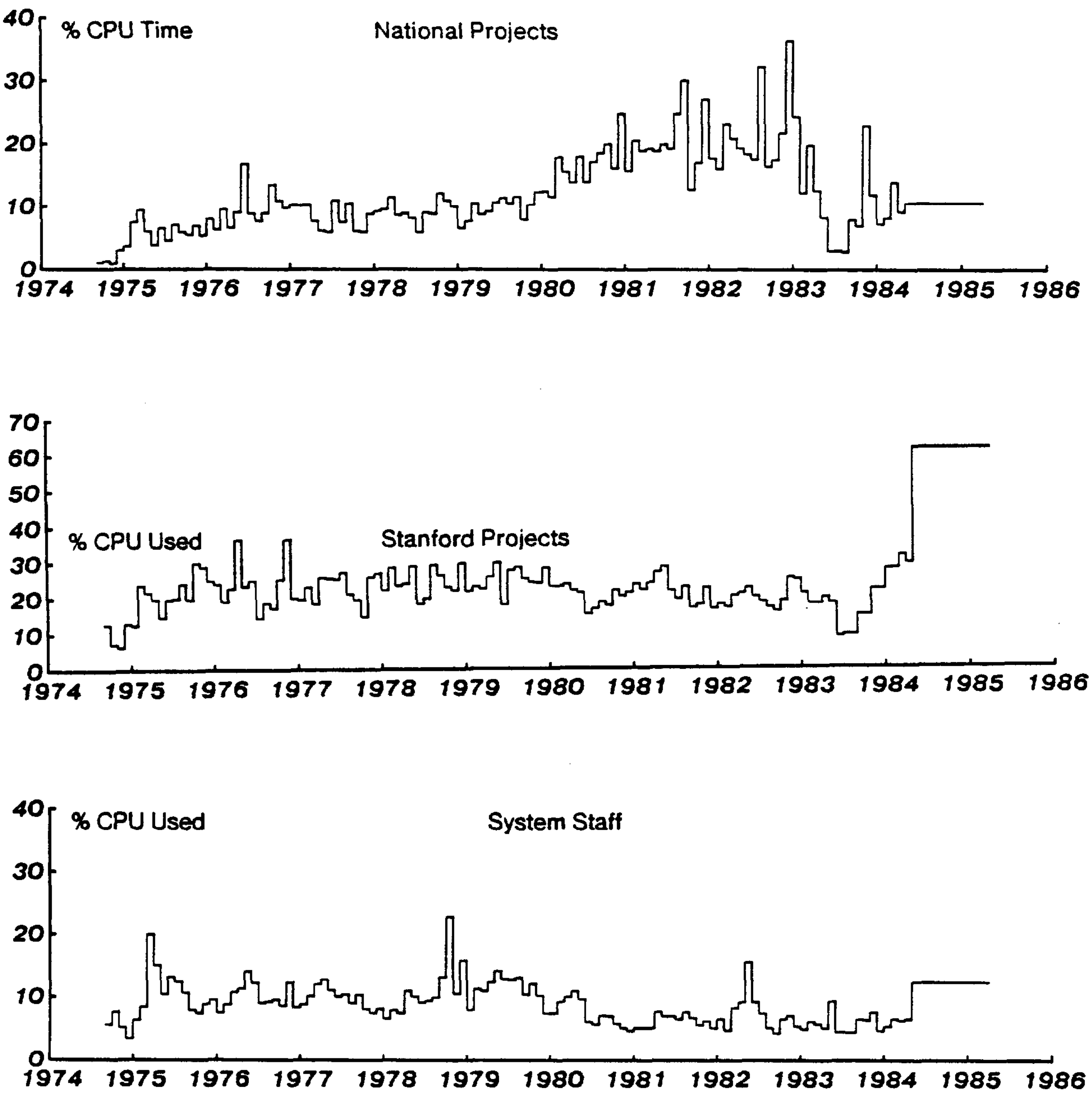

Figure 7: Monthly CPU Usage by Community

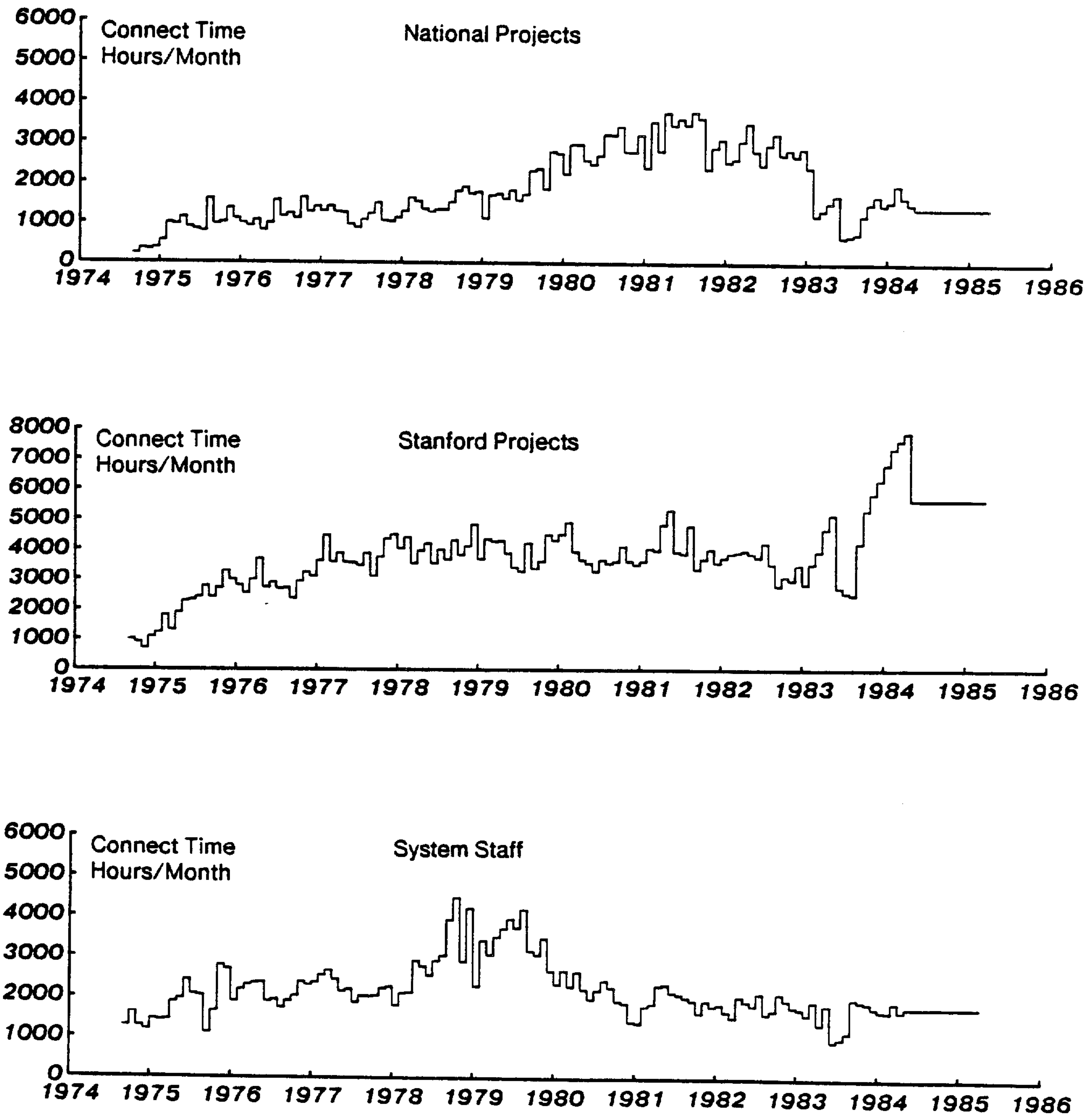

Figure 8: Monthly Terminal Connect Time by Community

Individual Project and Community Usage

The following histogram and table show cumulative resource usage by collaborative project and community during the past grant year. The histogram displays the project distribution of the total CPU time consumed between May 1, 1984 and April 30, 1985, on the SUMEX-AIM DECsystem2060 system.

In the table following, entries include a text summary of the funding sources (outside of SUMEX-supplied computing resources) for currently active projects, total CPU consumption by project (Hours), total terminal connect time by project (Hours), and average file space in use by project (Pages, 1 page = 512 computer words). These data were accumulated for each project for the months between May, 1984 and May, 1985.

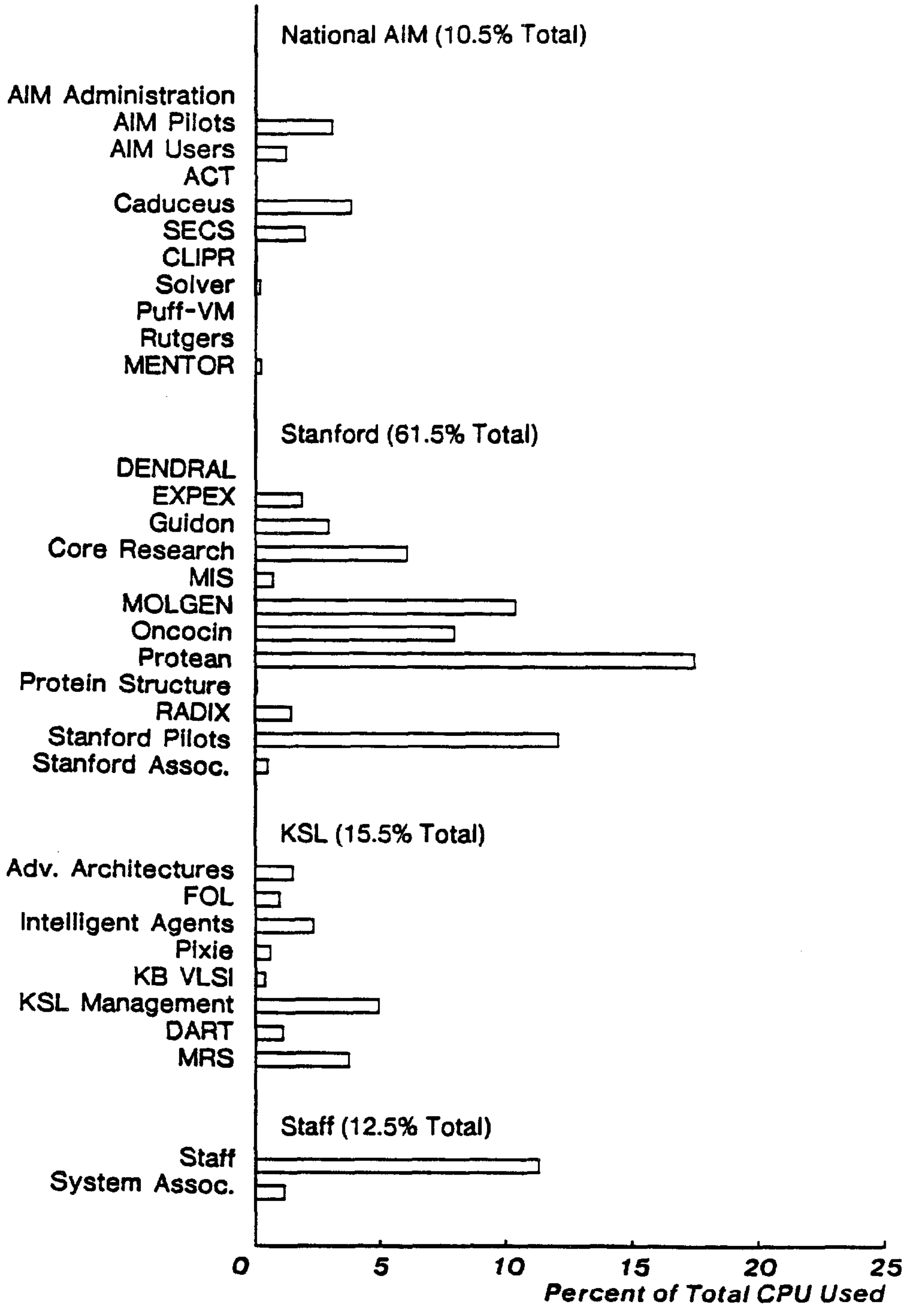

Figure 9: Cumulative CPU Usage Histogram by Project and Community

Resource Use by Individual Project - 5/84 through 4/85

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. University of Pittsburgh	86.72	1809.97	8028
2) CLIPR Project "Hierarchical Models of Human Cognition" Walter Kintsch, Ph.D. Peter G. Polson, Ph.D. University of Colorado	1.14	119.94	129
3) SECS Project "Simulation & Evaluation of Chemical Synthesis" W. Todd Wipke, Ph.D. U. California, Santa Cruz	45.14	5542.39	12230

4) SOLVER Project "Problem Solving Expertise" Paul E. Johnson, Ph.D. William B. Thompson, Ph.D. University of Minnesota	4.70	413.29	621
5) MENTOR Project "Medical Evaluation of Therapeutic Orders" Stuart M. Speedie, Ph.D. University of Maryland Terrence F. Blaschke, M.D. Stanford University	5.41	497.78	380
6) *** [Rutgers-AIM] *** Rutgers Research Resource Artificial Intelligence in Medicine Casimir Kulikowski, Ph.D. Sholom Weiss, Ph.D. Rutgers U., New Brunswick	0.62	57.29	196
7) AIM Pilot Projects	69.84	4292.54	3501
8) AIM Administration	0.42	57.86	673
9) AIM Users	27.88	3498.43	7135
	---------	----------	-------
Community Totals	241.87	16289.49	32893

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	67.60	8225.93	6048
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	238.64	8358.21	11392
3) ONCOCIN Project "Knowledge Engineering for Med. Consultation" Edward H. Shortliffe, M.D., Ph.D. Dept. Medicine	182.81	18869.06	16406

4) PROTEAN PROJECT Oleg Jardetzky School of Medicine Bruce Buchanan Computer Science Department	401.52	8539.01	13156
5) RADIX Project "Deriving Medical Knowledge from Time Oriented Clinical Databases" Robert L. Blum, M.D. Gio C.M. Wiederhold, Ph.D. Depts. Computer Science/ Medicine	33.23	2315.62	9168
6) Stanford Pilot Projects	277.71	6545.02	5092
7) Core AI Research	139.65	9447.97	10358
8) Stanford Associates	11.40	1030.22	1127
9) Medical Information Sciences	16.52	2561.42	974
	---------	----------	-------
Community Totals	1369.08	65892.46	70901

KSL-AI Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
For funding details please see page 47			
1) Advanced Architectures	34.45	11070.95	3313
2) FOL	22.61	781.19	1522
2) Intelligent Agent	53.25	6934.73	3205
3) Pixie	12.98	1989.63	1072
4) KB VLSI	8.47	1275.64	927
5) KSL Management	114.18	21341.80	15597
6) DART	25.05	1497.89	12677
7) MRS	86.40	9298.69	1950
	---------	----------	-------
Community totals	357.39	54190.52	40263

SUMEX Staff	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Staff	261.44	21450.55	17051
2) System Associates	26.84	1809.75	4744
	---------	----------	-------
Community Totals	288.28	23260.30	21795

System Operations	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Operations	775.69	69589.10	131640
	=====	=====	=====
Resource Totals	3032.31	229221.87	297492

(*) Award includes indirect costs.

System Reliability

System reliability for the DECsystem 2060 has significantly improved in this past period. We have had very few periods of particular hardware or software problems. The data below covers the period of May 1, 1984 to April 30, 1985. The actual downtime was rounded to the nearest hour.

Table 1 : System Downtime Hours per Month - May 1984 through April 1985

13	1	16	5	9	17	1	N/A	26	9	8	9
May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr

Table 2 : System Downtime Hours per Month - May 1984 through April 1985

Reporting period	:	364 days, 19 hours, 13 minutes, and 25 seconds
Total Up Time	:	359 days, 11 hours, 32 minutes, and 18 seconds
PM Downtime	:	1 days, 6 hours, 8 minutes, and 1 seconds
Actual Downtime	:	4 days, 1 hours, 33 minutes, and 6 seconds
Total Downtime	:	5 days, 7 hours, 41 minutes, and 7 seconds
Mtbf	:	3 days, 14 hours, 16 minutes, and 31 seconds
Uptime Percentage	:	98.89

Network Usage Statistics

The plots in Figure 10 and Figure 11 show the monthly network terminal connect time for the TYMNET and the INTERNET usage. The INTERNET is a broader term for what was previously referred to as Arpanet usage. Since many vendors now support the INTERNET protocols (IP/TCP) in addition to the Arpanet, which converted to IP/TCP in January of 1983, it is no longer possible to distinguish between Arpanet usage and Internet usage on our 2060 system.

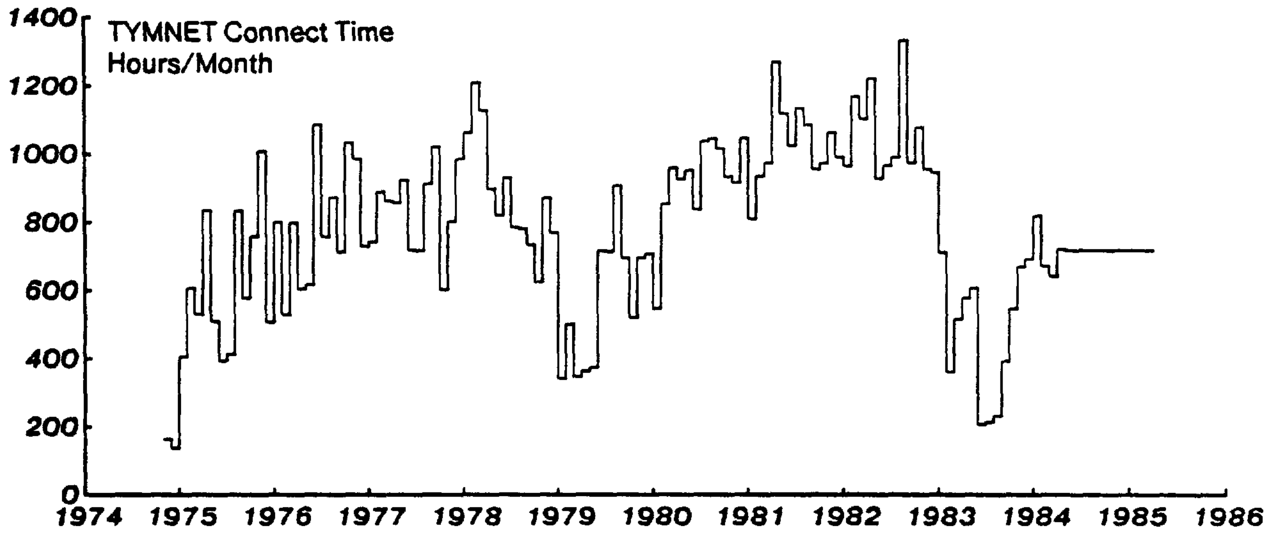

Figure 10: TYMNET Terminal Connect Time

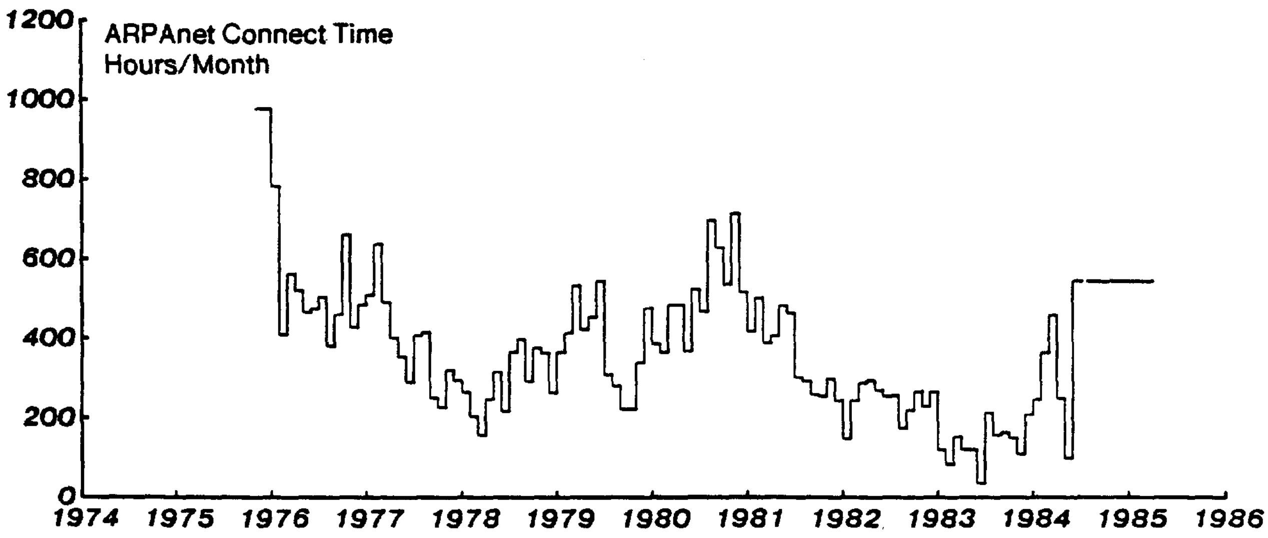

Figure 11: ARPANET Terminal Connect Time

III.A.4. Future Plans

Our plans for the next grant year (year 13) are based on the Council-approved plans for our 5-year renewal that began in August, 1980. The directions and background for much of this work were given in earlier progress report sections and are not repeated in detail here. Near- and long-term objectives and plans for individual collaborative projects are discussed in Section IV beginning on page 85.

Computing Resource Operation

The SUMEX-AIM resources -- mainframes, Lisp workstations, and networks -- provide crucial support for the AI research of our community. We will continue to operate these facilities for the most effective support of our users. We do not propose any substantial changes to the mainframe systems (DEC 2060, 2020, and shared VAX) but will continue to seek ways of minimizing maintenance costs and reliability.

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

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

Communication Networks

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

We have had reasonable success at meeting the geographical needs of the community through our ARPANET and TYMNET connections. These have allowed users from many locations within the United States and abroad to gain terminal access to the AIM resources and through ARPANET links to communicate much more voluminous file information. Since many of our users do not have ARPANET access privileges for technical or administrative reasons, a key problem impeding remote use has been the limited communications facilities (speed, file transfer, and terminal handling) offered currently by commercial networks. Commercial improvements are slow in coming and network delays have a major impact on remote projects -- mostly start-up pilot projects. We plan to continue experimenting with improved facilities as offered by commercial or government sources in the next year. We have budgeted for continued TYMNET service but will investigate alternatives as well, taking account of experiences with other national resources like BIONET.

Community Management

We plan to retain the current management structure that has worked so well in the past. We will continue to work closely with the management committees to recruit the additional high-quality projects which can be accommodated and to evolve resource allocation policies which appropriately reflect assigned priorities and project needs. We expect the Executive and Advisory Committees to play a continuing role in advising on priorities for facility evolution and on-going community development planning in addition to their recruitment efforts. The composition of the Executive Committee will continue to represent major user groups and medical and computer science applications areas. The Advisory Group membership spans both medical and computer science research expertise. We expect to maintain this policy.

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

The annual AIM workshops have served a valuable function in bringing community members and prospective users together. We will continue to support this effort. In July 1985, the AIM workshop will be hosted by the National Library of Medicine. We will continue to assist community participation and provide a computing base for workshop demonstrations and communications. We also will assist individual projects in organizing more specialized workshops as we have done for the DENDRAL and AGE projects.

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

Training and Education Plans

We have an on-going commitment, within the constraints of our staff size, to provide effective user assistance, to maintain high-quality documentation of the evolving software support on the SUMEX-AIM system, and to provide software help facilities such as the HELP and Bulletin Board systems. These latter aids are an effective way to assist resource users in keeping informed about system and community developments and solving usage problems. We plan to take an active role in encouraging the development and dissemination of community knowledge resources such as the AI Handbook, up-to-date bibliographic sources, and developing knowledge bases. Since much of our community is geographically remote from our machine, these on-line aids are indispensable for self-help. We will continue to provide on-line personal assistance to users within the capacity of available staff through the MM and TALK facilities.

Core Resource Development

Our primary focus for core resource development will be in the area of Lisp workstations including improvements to the computing environment they offer, facilitating their interaction with each other, and enhancing their interaction with network services. This will include bringing up tools like electronic mail, text processing, file management, and others that we currently relie almost entirely on mainframe computers for. We will study problems of high-performance network protocol and file service for these workstations as well as general access to network printing facilities. We will continue the development of virtual network and graphics interfaces for the workstations so they can be more geographically accessible and so their total computing power can be exploited.

Core AI Research

Our basic AI research projects focus on understanding the roles of knowledge in symbolic problem solving systems -- its representation in software and hardware, its use for inference, and its acquisition. We are continuing to develop new tools for system builders and to improve old ones. In particular, we will focus on four areas with immediate coupling to biomedical applications problems and on several others that may have future application. These include the Blackboard model of reasoning, constraint satisfaction systems, knowledge acquisition and learning, qualitative simulation, and other areas such as architectures for highly concurrent symbolic computation, a retrospective on the AGE blackboard tool, logic-based systems, self-aware systems, and the SOAR general problem-solving architecture.

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 some other special activities that reflect the impact and influence that the SUMEX-AIM resource has had in the scientific and educational communities.

III.B.1. Scholarly Publications

One of the important responsibilities of developers of a new technological area, such as artificial intelligence, is the scholarly assimilation and documentation of incremental progress. In addition to the numerous technical papers that have been published, 11 major books have been published from our community in the past 4 years:

- *Heuristic Reasoning about Uncertainty: An AI Approach*, Cohen, Pitman, 1985.
- *Readings in Medical Artificial Intelligence: The First Decade*, Clancey and Shortliffe, Addison-Wesley, 1984.
- *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Buchanan and Shortliffe, Addison-Wesley, 1984.
- *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*, Feigenbaum and McCorduck, Addison-Wesley, 1983.
- *Building Expert Systems*, F. Hayes-Roth, Waterman, and Lenat, eds., Addison-Wesley, 1983.
- *System Aids in Constructing Consultation Programs: EMYCIN*, van Melle, UMI Research Press, 1982.
- *Knowledge-Based Systems in Artificial Intelligence: AM and TEIRESIAS*, Davis and Lenat, McGraw-Hill, 1982.
- *The Handbook of Artificial Intelligence*, Volume I, Barr and Feigenbaum, eds., 1981; Volume II, Barr and Feigenbaum, eds., 1982; Volume III, Cohen and Feigenbaum, eds., 1982; Kaufmann.
- *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project*, Lindsay, Buchanan, Feigenbaum, and Lederberg, McGraw-Hill, 1980.

III.B.2. The PROTEAN Project

The biomedical goals of the PROTEAN project, under Professors Jardetzky and Buchanan at Stanford, are to use techniques from artificial intelligence to help in the determination of the 3-dimensional structure of proteins in solution. 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 diffraction 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 [16].

In the last year, the PROTEAN project has moved from the "idea phase" to the "demonstration phase":

- A highly interdisciplinary research team has been assembled which epitomizes the spirit of the SUMEX community. They include faculty in medicine and computer science, research associates in computer science (one with an MD degree), one MSTP graduate student, and other graduate students in Bio-Engineering, Chemical Engineering, and Computer Science (one with a PhD in Chemistry).
- A problem-solving framework, named BB1, has been debugged and is running. Much of the code already existed, some from the AGE system, but during this year it was extended and put into a coherent package. It is designed to be general enough to work with constraint satisfaction problems of many kinds, but has only been tested to date on the protein structure problem. One of the important extensions is to make reasoning about control as explicit as reasoning about objects in the domain.
- A geometry system has been designed and a prototype version has been written. This system is "low-level" code that manipulates objects in a 3-dimensional coordinate system and answers questions about locations, overlap, orientations, etc. This system depends on a representation of relative locations of objects with respect to an anchor, and with respect to one another. For example, if HELIX-1 is posted as the anchor, then HELIX-2 may be placed relative to HELIX-1 and other objects may be placed relative to HELIX-2.
- A manually operated version of PROTEAN was developed to allow Prof.Jardetzky and members of his laboratory to step through their own procedures for using NMR data to solve protein structures. The program allowed them to refine their procedures, and also allowed us to understand the procedures well enough to define knowledge sources that would carry out the same operations without manual intervention.
- A qualitative solution was found for the LAC-Repressor Headpiece, a protein of 51 amino acids. This approximate structure describes the relative positions of the three alpha-helices relative to one another, but does not place random coils. The structure is not completely determined by the constraints inferred from the NMR spectrum, so we have developed a representation of allowed volumes for the helices relative to one another.
- The IRIS graphics terminal has been coupled with the reasoning program to allow us to display the partial structures defined at any stage in the reasoning. The link is currently from a Xerox D-machine through a VAX to the IRIS. Display code, in the language C, has been written that allows

display of allowed volumes (as "halos" surrounding objects) and manipulation of objects on the screen.

- Knowledge sources have been written for BB1 that control the reasoning about solving protein structures. These define the heuristics used by biochemists to decide, for example, on the secondary structure to use as an anchor (the largest one with the most constraints with other parts of the structure). Enough knowledge sources have been defined so far to allow PROTEAN to reason autonomously through the first three-quarters of the problem solving cycles that biochemists use for the qualitative structure of the LAC-Repressor protein.

- A program, named MARCK, has been written that aids in the definition of new knowledge sources. It "watches" what an expert does manually to find the points at which the expert's reasoning and PROTEAN's reasoning diverge. Then it uses the problem solving context to help construct a new knowledge source that would make PROTEAN's reasoning agree with the expert's in that context. MARCK has been successfully used to define many of the control knowledge sources now in PROTEAN.

III.B.3. Software Export

The SUMEX-AIM community has widely distributed both our system software and our AI tool software to other academic, government, and industrial research groups in the United States and abroad. This form of "publication" allows others to critique our results and build on those foundations. To date, our AI tool exports include:

GENET	Prior to the establishment of the BIONET resource at IntelliCorp, we distributed 21 copies of the DNA sequence analysis programs and databases for both DEC-10 and DEC-20 systems.
EMYCIN	A total of 56 sites have received the EMYCIN [4, 34] package for backward-chained, rule-based AI systems.
AGE	The AGE [25] blackboard framework system has been sent out to 35 sites in versions for several machines.
MRS	The MRS [9] logic-based system for meta-level representation and reasoning has been provided to 76 sites.
Other Programs	Smaller numbers of copies of programs such as the SACON [2] knowledge base for EMYCIN, the GLISP [27] system (now distributed by Gordon Novak at the University of Texas), and the new BB1 [14, 13] system have been distributed.

A number of other software packages have been licensed or otherwise made available for commercial development including DENDRAL (to Molecular Designs, Ltd.), MAINSAIL (to Xidak, Inc.), UNITS (to IntelliCorp, Inc.), and EMYCIN (to Teknowledge, Inc. and Texas Instruments, Inc).

In addition, our system programs such as the TOPS-20 file recognition 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 well-distributed throughout the ARPANET community and the respective user communities.

III.B.4. The MENTOR Project

The MENTOR (Medical EvaluatioN of Therapeutic ORders) project is a transcontinental collaboration between Dr. Terry Blaschke at Stanford and Dr. Stuart Speedie at the University of Maryland. The MENTOR project was initiated in December 1983 as a pilot effort and has been funded by the National Center for Health Services Research since January 1, 1985. MENTOR's goal 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. Today, information is provided to the physician in the form of raw data which are 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, extensive medical knowledge and problem solving is required. An Expert System consultant on drug reactions could effectively gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base about drugs, 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.

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 use relatively simple criteria for possible reactions and do not try to represent the complex medical decision making process involved. 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.

The MENTOR system will use AI techniques to represent and reason about the complex of knowledge and data important to controlling adverse drug reactions in a monitoring and feedback system to influence physician decision-making. The initial effort has focused on the overall system design and work has begun on constructing a system for monitoring potassium in patients with drug therapy that can adversely affect potassium. Antibiotics, dosing in the presence of renal failure, and digoxin dosing have been identified as additional topics of interest.

III.B.5. Blackboard Model Research

Projects in the KSL, have experimented with many frameworks for building systems including rule-based (EMYCIN), frame-based (UNITS), logic-based (MRS), and blackboard-based (AGE) frameworks. We have also experimented with various methods of inference and control, including goal-directed, data-directed, and opportunistic reasoning. Of the paradigms we know about, the one that seems to offer the most flexibility is the blackboard model of reasoning.

It allows an arbitrary mixing of data-driven inference steps ("bottom up") with model-driven steps ("top down"). It allows a hierarchy of levels of abstraction in the on-going problem solution formation, from the most abstract (the global situation) to the least abstract (the supporting data or problem conditions). And it allows multiple sources of knowledge to provide the problem-solving links between these levels (i.e., information fusion).

Though the Blackboard framework was conceived at Carnegie-Mellon during the DARPA Speech Understanding project in the early 1970's, it has received much of its scientific and practical development by work in the Stanford Knowledge Systems Laboratory. The first development here was the HASP system for passive sonar signal understanding. Subsequent efforts involved experiments with scientific applications to x-ray crystallography, to planning, and in the development of the first software tool to assist knowledge engineers in constructing systems using the Blackboard framework (AGE).

The goal of our continuing research on blackboard systems is to improve the usability, the flexibility, and the inferential power of this framework for handling problems of hypothesis formation, signal understanding, constraint satisfaction and planning. This framework is also the organizing basis for our research on concurrent symbolic processing. We are implementing a new, domain-independent system called BB1 [13, 14] that incorporates a full range of blackboard tools and we are making these notions concrete by building a substantial application system in the BB1 framework in order to experiment with tradeoffs in the design. Specifically, BB1 is the basis for the PROTEAN project which is attempting to build a program that infers tertiary structure of proteins from NMR data (plus knowledge of primary and secondary structure).

BB1, like earlier blackboard systems, is a domain-independent "blackboard control architecture" that solves problems through the actions of independent knowledge sources that record, modify, and link individual solution elements in a structured database (the blackboard) under the control of a scheduler. It expands upon the standard architecture in that:

- It provides an interpretable, modifiable representation for knowledge sources with more flexible means for triggering appropriate ones and support facilities for knowledge source creation, modification, and checking.

- Its blackboard representation permits dynamic assignment of attributes and values to objects on the blackboard and provides selective, demand-driven inheritance of attributes from linked objects, with local caching of results.

- It provides explicit reasoning about control -- the selection and sequencing of knowledge source actions -- with control knowledge sources that construct dynamic control plans out of problem-solving heuristics on a control blackboard. It provides a vocabulary and syntax for expressing control heuristics and a simple scheduler decides which domain and control knowledge sources to execute by adapting to whatever control heuristics currently are recorded on the control blackboard.

- It provides strategic explanation of problem-solving activities.
- It provides generic learning knowledge sources to acquire new control heuristics automatically.
- Its run-time user interface provides capabilities for: displaying knowledge sources, pending actions, and objects on the blackboard; graphically displaying partial solutions via a user-specified interface; recommending pending actions for execution; permitting a user to override a recommendation; executing a designated action; and operating autonomously until a user-specified criterion is met.

BB1 is an evolving system that attempts to incorporate the best results of several research activities. We will continue developing BB1 as a prototype "next-generation" blackboard architecture.

III.C. Administrative Changes

Several administrative changes have occurred over the past year that affect the SUMEX-AIM resource.

In December 1984, the Knowledge Systems Laboratory (KSL) was formed as a reorganization of the Heuristic Programming Project (HPP -- see Appendix A). The new laboratory has a more modular organizational structure that recognizes the broad diversity of work now going on in the KSL and facilitates managing a research group of well over 100 people. The SUMEX-AIM resource continues to play a central role in KSL research.

On January 1, 1985, Mr. Edward Pattermann resigned as SUMEX Director to take a position involving AI tool development at IntelliCorp. He was replaced by Mr. Thomas Rindfleisch, who resumed the Director's role after two years managing the HPP. Mr. Rindfleisch retains his role (20%) as Director of the KSL and Mr. William Yeager has been appointed Assistant SUMEX Director to assist with day-to-day SUMEX management. Mr. Yeager has long been a key technical resource for SUMEX, having developed much of the Ethernet gateway and TIP service now in wide use.

Effective March 1, 1985, Dr. Edward Shortliffe was promoted to Associate Professor of Medicine with tenure. At the same time, Ted was appointed as Principal Investigator of SUMEX and Professor Feigenbaum resumed his role as co-Principal Investigator. This change in PIship in no way affects the long-standing interdisciplinary management of SUMEX, but just gave Dr. Shortliffe the appropriate title to carry out the key scientific and managerial role he already had been playing in SUMEX affairs. His active research has long been a core part of the SUMEX community and his Medical Computer Science group is physically co-located with SUMEX in the Stanford Medical Center. Also, SUMEX is located administratively in the Department of Medicine where Ted has his faculty appointment so he is in an excellent position to effectively represent the project with respect to its relationship with Stanford.

III.D. Resource Management and Allocation

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 B on page 211 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.

- We have allocated limited "collaborative linkage" funds as an aid to new projects or collaborators with existing projects to support terminals, communications costs, and other justified expenses to establish effective links to the SUMEX-AIM resource. Executive Committee advice is used to guide allocation of these funds.

- 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 have continued to provide active support for the AIM workshops. The last one was held at Ohio State University in the summer of 1984 and the next one will be in Washington, DC, hosted by the National Library of Medicine under Drs. Lindberg and Kingsland.

- We have continued our policy of no fee-for-service for projects using the SUMEX resource. This policy has effectively eliminated the serious administrative barriers that would have blocked our research goals of broader scientific collaborations and interchange on a national scale within the selected AIM community. In turn we have responded to the correspondingly greater responsibilities for careful selection of community projects of the highest scientific merit.

- 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.

As indicated above, we have sought to retain SUMEX resources for new projects, those exploring new areas in biomedical AI applications and those in such an early state of feasibility that they are unable to afford their own computing resources. This policy has worked effectively as seen from the following lists of terminated projects and projects now using their own computing resources at other sites:

Projects Moved All or In Part to Other Machines:

Stanford Projects:

- GENET [Brutlag, Kedes, Friedland - IntelliCorp]

National Projects:

- Acquisition of Cognitive Procedures (ACT) [Anderson - CMU]
- Chemical Synthesis [Wipke - UC Santa Cruz]
- Simulation of Cognitive Processes [Lesgold - Pittsburgh]
- PUFF [Osborne, Feigenbaum, Fagan - Pacific Medical Center]
- CADUCEUS/INTERNIST [Pople, Myers - Pittsburgh]
- Rutgers [Amarel, Kulikowski, Weiss - Rutgers]
- MDX [Chandrasekaran - Ohio State]
- SOLVER [P. Johnson - University of Minnesota]

Completed Projects Summary

Stanford Projects:

- DENDRAL [Lederberg, Djerassi, Buchanan, Feigenbaum]
- MYCIN [Shortliffe, Buchanan]
- EMYCIN [Shortliffe, Buchanan]
- CRYSALIS [Feigenbaum, Engelmore]
- MOLGEN I [Feigenbaum, Brutlag, Kedes, Friedland]
- AI Handbook [Feigenbaum, Barr, Cohen]

- AGE Development [Feigenbaum, Nii]

National Projects:

- Ventilator Management [Osborne, Feigenbaum, Fagan - Pacific Medical Center]
- Higher Mental Functions [Colby - USC]

III.E. Dissemination of Resource Information

Throughout the history of the SUMEX-AIM resource, we have made extensive efforts at disseminating the AI technology developed here. This has taken the form of many publications -- over 45 combined books and papers are published per year from 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 102).

Books and Publications

A sampling of the recent research paper publications of the KSL was given in the previous section on core AI research progress. The following lists the major books published in the past 4 years from the KSL:

- *Heuristic Reasoning about Uncertainty: An AI Approach*, Cohen, Pitman, 1985.
- *Readings in Medical Artificial Intelligence: The First Decade*, Clancey and Shortliffe, Addison-Wesley, 1984.
- *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Buchanan and Shortliffe, Addison-Wesley, 1984.
- *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*, Feigenbaum and McCorduck, Addison-Wesley, 1983.
- *Building Expert Systems*, F. Hayes-Roth, Waterman, and Lenat, eds., Addison-Wesley, 1983.
- *System Aids in Constructing Consultation Programs: EMYCIN*, van Melle, UMI Research Press, 1982.
- *Knowledge-Based Systems in Artificial Intelligence: AM and TEIRESIAS*, Davis and Lenat, McGraw-Hill, 1982.
- *The Handbook of Artificial Intelligence*, Volume I, Barr and Feigenbaum, eds., 1981; Volume II, Barr and Feigenbaum, eds., 1982; Volume III, Cohen and Feigenbaum, eds., 1982; Kaufmann.
- *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project*, Lindsay, Buchanan, Feigenbaum, and Lederberg, McGraw-Hill, 1980.

Software Distribution

We have widely distributed both our system software and our AI tool software. We have no accurate records of the extent of distribution of the system codes because their distribution is not centralized and controlled. The recent programs such as the TOPS-20 file recognition 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 well-distributed throughout the ARPANET community and beyond.

We do have reasonably accurate records of the distribution of our AI tool software because the recipient community is more directly coupled to us and the distribution is centralized:

GENET	Prior to the establishment of the BIONET resource at IntelliCorp, we distributed 21 copies of the DNA sequence analysis programs and databases for both DEC-10 and DEC-20 systems.
EMYCIN	A total of 56 sites have received the EMYCIN [4, 34] package for backward-chained, rule-based AI systems.
AGE	The AGE [25] blackboard framework system has been sent out to 35 sites in versions for several machines.
MRS	The MRS [9] logic-based system for meta-level representation and reasoning has been provided to 76 sites.
Other Programs	Smaller numbers of copies of programs such as the SACON [2] knowledge base for EMYCIN, the GLISP [27] system (now distributed by Gordon Novak at the University of Texas), and the new BB1 [14, 13] system have been distributed.

A number of other software packages have been licensed or otherwise made available for commercial development including DENDRAL (Molecular Designs), MAINSAIL (Xidak), UNITS (IntelliCorp), and EMYCIN (Teknowledge and Texas Instruments).

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.

The GENET Dissemination Experiment

Beginning in early 1980, the MOLGEN project investigators at Stanford have made a new set of computing tools available to a national community of molecular biologists through a guest facility called GENET on the SUMEX-AIM resource. This experimental subcommunity was started to broaden MOLGEN's base of scientist collaborators at institutions other than Stanford and to explore the idea of a SUMEX-like resource to disseminate sophisticated software tools to a generally computer-naive community. The enthusiastic response to the very limited announcement of this facility eventually necessitated SUMEX placing severe restrictions on the scope of services provided to this community.

Three main programs were offered to assist molecular genetics users: SEQ, a DNA-RNA sequence analysis program; MAP, a program that assists in the construction of restriction maps from restriction enzyme digest data; and MAPPER, a simplified and somewhat more efficient version of the MOLGEN MAP program, written and maintained by William Pearson of Johns Hopkins University. Some of the other, more-sophisticated programs being developed through MOLGEN research efforts were not yet available for novice users. However, GENET users had access to the SUMEX-AIM programs for electronic messaging, text-editing, file-searching, etc.

The GENET experiment proved so successful that eventually that community was the single biggest consumer of processor cycles on SUMEX. This overload diverted our very limited computing resources away from our mainline goal of supporting projects developing new AI systems in the medical and biological sciences, including molecular biology. Efforts to secure funds to increase SUMEX capacity for the burgeoning GENET use failed. Thus, without any fair way to allocate a small resource to the growing GENET community and in order to restore the necessary emphasis on biomedical computer science research on SUMEX, it was necessary to phase out the GENET usage. We closed the GENET account at the end of 1982, with a mandate from an ad hoc GENET Executive Committee, and phased out all usage by spring of 1983. In the process, we developed procedures by which academic users could obtain their own copies of the GENET programs used at SUMEX and we provided a list of alternate sources for GENET-like computing services. As indicated above, SUMEX has supplied 21 systems to academic users with compatible machines.

Since the phase-out of GENET at SUMEX, IntelliCorp, a commercial AI company, submitted a proposal to the NIH Division of Research Resources for a BIONET resource and was successful in obtaining funding. The BIONET resource began operation in the summer of 1984.

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 C on page 215.

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. This work derived from our first system, the MYCIN program. That research led to three sub-projects (EMYCIN, GUIDON, and ONCOCIN) described in previous annual reports. EMYCIN has been completed and its resources reallocated to other projects. GUIDON and ONCOCIN have become projects in their own right.

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

The purpose of the current contracts is to construct an intelligent tutoring system that teaches diagnostic strategies explicitly. By strategy, we mean plans for establishing a set of possible diagnoses, focusing on and confirming individual diagnoses, gathering data, and processing new data. The tutorial program will have capabilities to recognize these plans, as well as to articulate strategies in explanations about how to do diagnosis. The strategies represented in the program, modeling techniques, and explanation techniques are wholly separate from the knowledge base, so 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 utilized in additional ways that are medically relevant. Using the knowledge to teach medical students is perhaps foremost among these, and NEOMYCIN continues to focus on methods for augmenting clinical knowledge in order to facilitate its use in a tutorial setting. A particularly important aspect of this work is the insight that has been gained regarding the need to structure knowledge differently, and in more detail, when it is being used for different purposes (e.g., teaching as opposed to clinical decision making). It was this aspect of the

GUIDON research that led to the development of NEOMYCIN, which is an evolving computational model of medical diagnostic reasoning that we hope will enable us to better understand and teach diagnosis to students. An important additional realization is that these structuring methods are beneficial for improving the problem-solving performance of consultation programs, providing more detailed and abstract explanations to consultation users, and making knowledge bases easier to maintain.

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

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

C.1.1 The NEOMYCIN Consultation Program

NEOMYCIN is distinguished from other AI consultation programs by its use of an explicit set of domain-independent metarules for controlling all reasoning. These rules constitute the diagnostic procedure that we want to teach to students: the stages of diagnosis, how to focus on new hypotheses, and how to evaluate hypotheses. This diagnostic procedure as well as the knowledge base underlying the procedure has remained relatively stable this year. Our work in explanation highlighted the importance of making the knowledge used by the system at all levels as explicit as possible. As a result, this year we have extended and refined a previous predicate calculus representation of NEOMYCIN's metaleval rules. To avoid earlier problems of efficiency with this representation, we have also written a compiler that produces Lisp code from our predicate calculus notation. As a result, we are able to run the more efficient Lisp code and use the explicit notation for explanation and modeling.

To develop and test our model of heuristic classification, we are producing from NEOMYCIN a generic system, called HERACLES, that can be used to solve other problems by classification. This is an "E-NEOMYCIN," NEOMYCIN without its current medical knowledge. HERACLES is a variant of EMYCIN; it enables a knowledge engineer to produce NEOMYCIN-like knowledge bases containing the NEOMYCIN diagnostic procedure and domain knowledge organization. To prove its true generality, our first HERACLES knowledge base is in the manufacturing domain, for diagnosing sand casting problems (for the process of forming metal objects using sand molds). Future knowledge bases could be drawn from many medical and non-medical domains.

C.1.2 The ODYSSEUS Modeling System

This effort concerns automation of the transfer of expertise between an expert system and a human expert. A major goal is to produce a system that can watch an expert solve a problem and automatically recognize *differences* between the expert's underlying knowledge base and an expert system's knowledge base. This system should demonstrate how a knowledge of these differences can aid knowledge acquisition and intelligent tutoring. The program implementing this approach, called ODYSSEUS, has several stages of operation. Based on a large set of problem-solving sessions, the program first induces the rule and frame knowledge to drive HERACLES. Using this initial knowledge base as a "half-order theory," subsequent problem-solving sessions are tracked step by step: for each observable step the specialist makes, ODYSSEUS generates and scores the alternative *lines of reasoning* that can explain the specialist's reasoning step. When no plausible reasoning path is found, or all found ones have a low score, the program assumes it is deficient in either its strategic or domain knowledge. It attempts to acquire the missing knowledge either automatically or by asking the specialist specific questions. In a variation, the specialist justifies each problem-solving step using the vocabulary of an abstract justification language. These justifications aid in scoring alternative plausible lines of reasoning.

Each of the stages of ODYSSEUS has been implemented as a separate subsystem. These subsystems are now being integrated.

C.1.3 The NEOMYCIN Explanation System

The initial explanation system of NEOMYCIN enables the user to ask WHY and HOW questions during a consultation. That is, when the program prompts the user for new data, the user may ask WHY the data is being requested or HOW some strategic task will be (or was) accomplished. Unlike MYCIN's explanation system, upon which this kind of capability is patterned, explanations in NEOMYCIN are in terms of the diagnostic plan, not just specific associations between data and diagnoses.

The next phase of this work is to answer WHY questions by condensing the entire line of reasoning. The program uses general explanation heuristics, models of the user's knowledge of diseases and of strategy, and a history of the user's interaction with the current consultation to select the task, focus, and domain information that is most likely to be of interest. Some of the heuristics used by the explanation system include: 1) mentioning the last task whose focus (or argument) changed in kind (e.g. from a disease hypothesis to a finding request); 2) never mentioning tasks that are merely iterating over a list of rules, findings, or hypotheses; and 3) only mentioning tasks with rules as an argument to programmers. These heuristics, as well as the general procedure for providing explanations, have been implemented in the same task and metarule language used to represent NEOMYCIN's diagnostic strategy. In addition, the explanation system has been extended to use the MRS version of the task metarules. We are thus able to select the specific medical relations that were used by the metarule in determining what action to take. As a result, we have more detailed and concise information to explain to the user. The clearer representation of both the information that can be explained and the explanation procedure provides us with a flexible, explicit encoding of our method for producing explanations, which will serve as a basis for devising tutoring techniques, as well as understanding explanations provided by users of their diagnostic strategy.

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

C.1.4 Graphics for Teaching

We are continuing to make extensive use of graphics in our programs. As part of our series of instructional programs, GUIDON-WATCH has been implemented as a graphics system for watching NEOMYCIN's reasoning. For example, we can highlight the hypothesis under consideration in the diagnostic taxonomy and show graphically how the program "looks up" its hierarchies before refining hypotheses. In addition, the user is able to explore the findings, hypotheses, rules and tasks that comprise the knowledge base, see selected causal association networks, view the differential as it changes, and keep track of hypotheses with evidence and positive findings. All of these can be easily selected with a consistent menu system, and windows on the screen are automatically organized to clearly display the information requested by the user.

C.2 Research in Progress

The following projects are active as of June 1984 (see also near-term plans listed in Section III.A):

1. development of a prototype of a bottom-up student modeler

2. standardization of display code

3. prototype of GUIDON-MANAGE

4. prototype of HERACLES and demonstration in non-medical domain

5. user model incorporated in explanations, with summarization

6. student model learning discrepant domain knowledge

D. Publications Since January 1984

1. Clancey, W. J.: *Knowledge acquisition for classification expert systems.* Proc. ACM-84. Also Heuristic Programming Project Report HPP 84-18, Computer Science Dept., Stanford Univ., July, 1984.

2. Clancey, W.J.:*Heuristic classification.* Knowledge Systems Laboratory Report KSL 85-5, Computer Science Dept., Stanford University, March 1985.

3. Richer, M., and Clancey, W.J.: *GUIDON-WATCH: A graphic interface for browsing and viewing a knowledge-based system.* Submitted to IEEE.

4. Wilkins, D.C., Buchanan, B.G., and Clancey, W.J.: *Inferring an expert's reasoning by watching.* Proc. 1984 Conference on Intelligent Systems and Machines, Rochester, MI, April 1984, pp.51-58.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

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

B. Sharing and Interaction with Other SUMEX-AIM Projects

We plan to add learning capabilities of two forms into this framework, involving interactions with the machine learning group within the KSL and Prof. Paul Rosenbloom's project on SOAR.

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

C. Critique of Resource Management

The SUMEX staff has been extremely helpful in maintaining connections between Xerox D-machines and SUMEX. The SUMEX staff also rewrote communication software used to link the D-machines to SAFE, the file saver used by the GUIDON/NEOMYCIN group. This has greatly improved both performance and reliability.

III. RESEARCH PLANS

A. Project Goals and Plans

Research over the next year will continue on several fronts, leading to several prototype instructional programs by early 1986.

1. Test student modeling program on cases chosen for teaching, collecting data for further development of the program, as well as exploring the range of student approaches to diagnosis.

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

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

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

5. Refine NEOMYCIN diagnostic model (relations and procedures) by student modeling and knowledge acquisition efforts.

6. Develop, debug, and document an exportable version of HERACLES, a generic knowledge engineering tool that can be used to produce additional medical and non-medical knowledge bases to be tutored by GUIDON2.

7. Formalize heuristics for teaching, given the NEOMYCIN model and heuristics for explanation and modeling, embodied in different versions of GUIDON2.

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

We sketch here our general conception of the research we plan for 1985-88, specifically the GUIDON2 family of instructional programs, based on the NEOMYCIN problem-solving model. Our ideas are strongly based on recent proposals by J.S. Brown, particularly his paper "Process versus Product -- A perspective on tools for communal and informal electronic learning" and some related papers that he wrote in 1983, in which he proposes methods for giving a student the ability to reflect on how he solved a problem. We have designed a family of seven programs that as a sequence will teach students to think about their own thinking process and to adopt efficient, effective approaches to medical diagnosis.

The key idea is that NEOMYCIN provides a *language* by which a program can converse with a student about strategies and knowledge organization for diagnosis. NEOMYCIN's tasks and structural terms provide the *vocabulary* or *parts of speech*; the meta-rules are the *grammar* of the diagnostic process. We will construct different graphic, reactive environments in which the student can observe, describe, compare, and improve his own diagnostic behavior and that of others. By "reactive environment" we mean that these programs are not passive, they will watch what the student does, build a model of his understanding and learning preferences, and provide corrective advice.

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

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

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

GUIDON-ANNOTATE This system allows the student to annotate a NEOMYCIN typescript, *explaining* in strategic and/or domain terms what the program is doing each time it requests new case data, indicating the task and focus associated with each data request. The program will indicate, upon request, where the student is incorrect and which annotations are different from NEOMYCIN's, but are still reasonable interpretations. The student will be able to choose these tasks from a menu of icons, either linearly or hierarchically displayed, as he prefers. (Again, NEOMYCIN will annotate its own solutions upon request and allow replaying.) This activity gets the student to think strategically by recognizing a good strategy. In this way, he learns to recognize how strategies affect the problem space.

GUIDON-APPRENTICE This is a variant of NEOMYCIN in which the program stops during a consultation and asks the student to propose the next data request(s). The student is asked to indicate the task and focus he has in mind, plus the differential he is operating upon. The program compares this proposal to what NEOMYCIN would do. In this activity we descend to the domain level and require the student to instantiate a strategy appropriately. Ultimately, such a program will use a *learning model* that anticipates what the student is ready to learn next and how he should be challenged. Early versions can simply use built-in breakpoints supplied by an expert teacher. In the future, programs will develop their own curriculums from a case library.

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

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

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

C. Requirements for Continued SUMEX Use

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

D. Requirements for Additional Computing Resources

With the addition of two new D-machines for this work, our computing needs will be adequately met in the coming 1-2 years at least.

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

E. Recommendations for Future Community and Resource Development

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

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

From this perspective, future resource development should focus on improving the capabilities of networks, file servers, and attached devices to respond to individual requests. Multi-processing becomes a necessity in such an environment, so a request can be honored while the user returns to continue his programming or editing.

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 200 new and established industrial firms already undertaking product development using these new genetic tools.

The programs being developed in the MOLGEN project have already proven useful and important to a considerable number of molecular biologists. Currently several dozen researchers in various laboratories at Stanford (Prof. Paul Berg's, Prof. Stanley Cohen's, Prof. Laurence Kedes', Prof. Douglas Brutlag's, Prof. Henry Kaplan's, and Prof. Douglas Wallace's) and over 400 others throughout the country have used MOLGEN programs over the SUMEX-AIM facility. We have exported some of our programs to users outside the range of our computer network (University of Geneva [Switzerland], Imperial Cancer Research Fund [England], and European Molecular Biology Institute [Heidelberg] are examples). The pioneering work on SUMEX has led to the establishment of a separate NIH-supported facility, BIONET, to serve the academic molecular biology research community with MOLGEN-like software. BIONET is now serving many of the computational needs of over 1000 academic molecular biologists in the United States.

C. Highlights of Research Progress

C.1 Accomplishments

The current year has seen the completion of our initial study of the Yanofsky project on genetic regulation in the *trp* operon. In addition we have tested several models of qualitative simulation of biological systems and begun our design of a theory discovery system. Finally, a new application program for DNA sequence analysis was developed by one of our research collaborators. The highlights of this work are summarized in several categories below.

C.1.1 The Scientific Process of Theory Formation, Modification, and Testing

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

We have conducted extensive interviews with Professor Yanofsky and many of his former students and collaborators. We have examined most of the relevant research papers. We believe we now have a good understanding of the three major classes of knowledge that were important in the discovery of the theory of regulation in the trp operon: knowledge about the relevant biological objects, knowledge about the techniques used to elicit new information, and discovery heuristics used to build new models.

In addition, we have developed an initial model for the inference mechanisms used during the discovery process. This model includes at least four different types of reasoning: data-driven, theory-driven, analogy to closely-related biological systems, and analogy to other systems (railroad engines and tracks, for example).

C.1.2 Knowledge-Based Simulation of the Trp Operon

The first major programming task of our project was to build a knowledge base representing the initial state of knowledge about the tryptophan operon system at the beginning of the Yanofsky research. This initial knowledge base contains information relevant to genetic regulation in general and to the trp operon system in particular. The information relates both to structure, i.e. the physical characteristics of the biological objects, and to function, i.e. the operational characteristics of the biological objects. In addition, the procedural knowledge needed to relate structure to function plays an important part in the knowledge base.

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

As a first experiment, we built the knowledge base using the Unit System (developed under previous MOLGEN work). We were able to successfully model most of the important processes of Jacob-Monod repression, the initial model of genetic regulation used in the Yanofsky research.

C.1.3 A Model for Theory Discovery

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

In addition, we have fit our overall model of simulation and discovery into a framework of research on machine learning. This framework involves interacting performance and learning elements. The performance element, here the knowledge-based system for qualitative simulation of regulatory genetics, is asked to explain observations from the real world. The learning element, here the discovery architecture described above, is able to evaluate the explanations and "tune" the performance element by changing its model (or theory) of the world.

C.1.1.4 Simultaneous alignment of DNA sequences--MULTAN

Previously, MOLGEN researchers have developed numerous programs to aid in the symbolic analysis of DNA sequences. During the last year Dr. William Bains (a postdoctoral scholar in Professor Kedes' laboratory), completed a program called MULTAN which allows the facile alignment of three or more DNA sequences. This was a major unsolved problem in sequence analysis and the program is now undergoing final testing on the BIONET resource. In the future, we expect that BIONET will support development of application-oriented programs of this type, while MOLGEN and SUMEX will focus on research-oriented systems with major AI goals.

C.2 Research in Progress

We have two major goals over the next several months. The first is to convert and enhance our knowledge-based simulation model within the KEE tool from IntelliCorp, Inc. KEE will be a significant improvement over the Unit System in three areas: speed, functionality, and support. IntelliCorp is providing KEE for use in our research without charge. Studies have indicated that using KEE will unable us to produce a reasonable prototype of our discovery system in about half the time or using the Unit System. Our second goal is to more formally define the learning element of our discovery system and to build a first test system that operates upon the simulation system knowledge base.

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*. (To appear in a joint issue of Communications of the ACM and IEEE/Computer, ~~October,~~ 1985). CACM 28(11): 1164-1186

12. Friedland, P. and Iwasaki Y.: *The concept and implementation of skeletal plans*. (To appear in Journal of Automated Reasoning, Vol. 1, No. 2, 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.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

SUMEX-AIM continues to provide the bulk of our computing resources. The facility has not only provided excellent support for our programming efforts but has served as a major communication link among members of the project. Systems available on SUMEX-AIM such as INTERLISP, TV-EDIT, and BULLETIN BOARD have made possible the project's programming, documentation and communication efforts. The interactive environment of the facility is especially important in this type of project development.

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

We have taken advantage of the collective expertise on medically-oriented knowledge-based systems of the other SUMEX-AIM projects. In addition to especially close ties with other projects at Stanford, we have greatly benefited by interaction with other projects at yearly meetings and through exchange of working papers and ideas over the system.

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

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

III. RESEARCH PLANS

A. Project Goals And Plans

Our current work has the following major goals:

1. Use the knowledge base to explain observations that are indeed explainable without changes to the current model. For example, "I have observed a mutation that causes constitutive (uncontrolled) production of tryptophan. How can that be explained within the Jacob-Monod model?" This process will be accomplished by some combination of forward simulation and backward rule-chaining.

2. Begin to recognize when observations are "interesting." Interesting here has one of the following broad meanings:

   a. A seeming direct contradiction to the existing theory.

   b. A statistically rare occurrence (one that is understandable by the current theory, but should not occur very often).

   c. A dramatic confirmation of the existing model.

   d. An observation currently unpredictable by the current model because the model is either not detailed enough or incomplete. The observation in this case must have a relation to the model because an important object of the model is involved or it relates to an effect predicted by the model.

3. Build a mechanism for postulating extensions or corrections to the current theory: a contrained regulatory theory generator. The overall approach to this mechanism is perhaps the most interesting problem in our work. In discussions with other computer scientists, the notion of "or" reasoning where the theory construction process consists of hierarchical refinement of abstract ideas into more detailed ones, and "and" reasoning where the theory is built up in little pieces at many different levels simultaneously has emerged. We see strong evidence for both types of reasoning within Yanofsky's project. In fact, as stated above, the global model of Yanofsky's laboratory is a hybrid one. Individual graduate students performed "and" tasks--filling in details of seemingly unrelated pieces of the model. Yanofsky was the master "or" reasoner, slowly building a hierarchical model of the new regulatory mechanism. It is in this area of our research where the greatest discussion with AI colleagues is needed and which may produce the most significant AI benefits.

4. 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.

5. 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, etc.

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.

Responses to Questions Regarding Resource Future

1. role of SUMEX after 7/86--I strongly believe that the 2060 should have continuing support for the forseeable future. The maturity of software for communications, document preparation, and general support of scientific literacy is unsurpassed. One has only to note the heavy continued load on SUMEX, despite the proliferation of workstations, VAXes, etc. around the KSL to see that it is still being used productively. In addition, the ability to easily work from home at all hours contributes greatly to overall productivity within the SUMEX community.

2. will my group require continued access--Yes, very much so for all of the reasons outlined above.

3. impact of user fees--Modest user fees would not have an enormous impact, but would prevent the kind of easy, productive use for general purposes that SUMEX now serves. I think the greater impact would be on not fully established or new research groups during start-up mode.

4. workstation plans--my group, MOLGEN, already makes extensive use of workstations for mainline computing purposes. Despite this use, we still find the SUMEX 2060 invaluable.

I would add to #1, that continuing research on melding together a distributed environment, of which both single-user workstations and the 2060 are parts should be a major continuing goal of SUMEX research.

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 insure that the system is easy to use and acceptable to physicians. The work seeks to improve the interactive process, both for the developer of a knowledge-based system, and for the intended end user. In addition, we have emphasized clinical implementation of the developing tool so that we can ascertain the effectiveness of the program's interactive capabilities when it is used by physicians who are caring for patients and are uninvolved in the computer-based research activity.

B. Medical Relevance and Collaboration

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

C. Highlights of Research Progress

C.1 Background and Overview of Accomplishments

The ONCOCIN Project is a large interdisciplinary effort that has involved over 35 individuals since the project's inception in July 1979. With the work currently in its sixth 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 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 have 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.

With daily coordination by the project's data manager, Janice Rohn, the DEC-20 version of ONCOCIN continues to be used on a limited basis in the Stanford Oncology Clinic. The continued dependence on this time-shared computer, however, has prevented us from using ONCOCIN in in many clinical problem areas (other than the lymphomas where clinics are held three mornings per week, and breast cancer where clinic is held one day per week) because of our inability to assure the system's availability with reasonable response time. It is this latter point that has accounted for our decision not to spend a great deal of time developing new protocols to run on the DEC-20 ONCOCIN prototype. Instead we have pressed our effort to adapt ONCOCIN to run on professional workstations which can eventually be dedicated to full time clinic use. We envision these workstations as the model for eventual dissemination of this kind of technology.

In addition to funding from DRR for the workstation conversion effort, we have support from the National Library of Medicine that supports our more basic research activities regarding biomedical knowledge representation, knowledge acquisition, therapy planning, and explanation as it relates to the ONCOCIN task domain. A grant from the NLM to study the therapy planning process was received, and this work (led by Dr. Fagan) is in its second year. This research is investigating how to represent the therapy planning strategies used to decide treatment for patients on the oat cell carcinoma protocol who run into serious problems requiring consultation with the protocol study chairman. Dr. Branimar Sikic, a faculty member from the Stanford University Department of Medicine, and the Study Chairman for the oat cell protocol, is collaborating on this project.

C.2 Research in Progress

The major efforts of the ONCOCIN project over the last year have fallen into three major categories: (1) conversion of ONCOCIN to run on workstations, (2) development of a knowledge acquisition interface (OPAL) for entering new protocols, and (3) research on modeling the strategic therapy selection process (ONYX). Efforts are also in progress to evaluate the system, to document the results of the research, and to disseminate the technology to sites beyond Stanford. We summarize these ongoing research efforts below.

C.2.1 Transfer of the ONCOCIN system from the DEC-20 to the Xerox 1108

In an effort to improve the efficiency of the reimplemented system (and thereby to improve its response time and make it more acceptable to physicians), we have undertaken a substantial system redesign while transferring it to the new machines. An additional commitment in time and programming effort has resulted, but we are confident that the resulting system will be a substantial improvement over the prototype. There have been several aspects to the system's reimplementation during the current year:

- *Reorganization and recoding of existing programs for improved efficiency.* In last year's report, we discussed our first steps in reorganizing the program. A further analysis during the year suggested that we should consider a redesign of the program to take advantage of our experience with the existing program and to respond to advances in artificial intelligence representation methods since ONCOCIN was first designed. In addition, our work during the year on new methods for entering knowledge into the system suggested corresponding improvements in the ways to represent oncologic knowledge in the system (see paper by Musen, et al. for more details on the redesign of the ONCOCIN system).

- *Redesign of the reasoning component.* As a major part of the redesign of the system, we decided to concentrate on methods that would allow for a more efficient search of the knowledge base during the running of a case. We have implemented and are currently debugging a reasoning program that uses a discrimination network to process the cancer protocols. This network allows for a compact representation of information that overlaps elements of multiple protocols, but does not require the program to consider and then disregard information related to protocols that are irrelevant to a particular patient.

- *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 (see the paper by Kahn, et al. for a more detailed description of the temporal network).

- *Extensions to the user interface.* The user interface has been extended so that it can read patient data files of the type that are created by the original ONCOCIN system. This will allow us to transfer currently active patients to the new version of the ONCOCIN system. A detailed description of the user interface is available in the paper by Lane, et al.

- *Connecting the components of the ONCOCIN system.* The reasoning component, user interface, and knowledge acquisition program (described below) have been developed as separate programs. In the final version of the system, the knowledge acquisition program must be able to automatically translate from the graphical input forms into the knowledge base. The reasoner and user interface components are independent programs that run in parallel while communicating with each other. Each of these connections between components has been tested on a limited basis and will continue to be exercised during the next several months.

- *Knowledge engineering tools.* The challenge of coordinating a large software development project, with multiple programmers working in parallel, has necessitated the development of specialized tools to facilitate the process of system construction and maintenance. One area of particular concern has been the need for tools to assist with knowledge base maintenance (see paper by Tsuji and Shortliffe for a discussion of our initial work in this area).

- *System support for the reorganization.* The LISP language that we used to build the first version of ONCOCIN does not explicitly support basic knowledge manipulation techniques (viz. 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 provides the reliability, speed, size, or special memory-manipulation techniques that are needed for our project.

We have accordingly developed a "minimal" object-oriented system to meet these specifications. The object system is currently in use by each component of the new version of ONCOCIN and in the software used to connect the components. In addition, several student projects are now able to use this programming environment.

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

A major effort in this grant year has been the development of software (termed the OPAL system) that will permit physicians who are not computer programmers to enter protocol information into a structured set of forms on a graphical display. Most early expert systems required tedious (and occasionally erroneous) entry of the system's medical knowledge. Each segment of knowledge was transferred from physician to programmer and then entered into the program by the computer expert. Although many programs allowed for specification of a structure within which to organize the information, only minimal attempts were made to define a description that would be generic enough to provide a basis for a series of related knowledge bases in one medical area.

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: multiple protocols (that may be related to each other), experimental research arms in each protocol, drug combinations, individual drugs, and drug modifications. Using the graphically-oriented workstations, this information is presented to the user as computer-generated forms that appear on the screen. As the protocol is described, new forms are added to the computer display to allow for the specification of the special cases that make the protocols so complicated.

Although this design appears to be organized specifically for cancer treatment plans, we believe that the technique 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 could accept *any* possible problem specification.

Using this program we have entered several versions of a small cell lung cancer protocol, and a complicated lymphoma protocol with several different therapies. We are currently implementing the changes suggested by entering these protocols.

C.2.3 Strategic Therapy Planning (ONYX)

As mentioned above, we have begun a new research project to study the therapy planning process, and how strategies which are used to plan therapy in difficult cases might be represented on a computer. This project, which we call the ONYX project, has as its goals: to conduct basic research into the possible representations of the therapy planning process; to develop a computer program to represent this process; and eventually to interface the planning program with ONCOCIN. The project members (Fagan, Tu, Langlotz, and Williams) have spent many hours meeting with Dr. Sikic trying to understand how he plans therapy for patients whose special clinical situation precludes following the standard therapeutic plan described in the protocol document. In March of last year, the group spent two days at Xerox Palo Alto Research Center (PARC), working with Mark Stefik, Daniel Bobrow and Sanjay Mittal of PARC on possible representations for the knowledge structures and how such a program might run using the LOOPS knowledge programming system. A prototype version of this program is currently being tested. The prototype program has been designed as two components: the strategic planning program and the qualitative simulation builder. The strategic planning program is capable of turning the patient's medical data and knowledge of the intent of the protocol into a small number of plausible protocol modifications for the current point in time, and conditional modifications for the near future. Another component of the system is capable of building simulation models using the graphical abilities of the 1108 workstation. The first test of this component is the construction of a model of the effects of chemotherapy drugs on the bone marrow of the patient. During the next year of research this type of qualitative simulation model will be integrated into the strategic planning program.

C.2.4 Evaluations of ONCOCIN's performance

We have completed our first three formal studies of ONCOCIN's DEC-20 version (see papers by Kent et al. and Hickam et al. for results of two of these; written reports on the third is in preparation). Lessons learned in these initial studies have led to revisions both in the design of ONCOCIN and in our plans for evaluation studies of the 1108 version of the system when it is implemented at non-Stanford sites in later years.

C.2.5 Documentation

We have developed a videotape that discusses and demonstrates our research on the workstation version of our system. This tape has been shown at national meetings and has been extensively distributed to researchers internationally who have shown an interest in our work. The publication list that accompanies this report further documents the design decisions we have made in developing the new version of ONCOCIN.

C.2.6 Dissemination

In anticipation of completion of the workstation version of ONCOCIN, we are beginning to plan for an experiment in which we will install ONCOCIN workstations in private oncology offices in San Jose and Fresno. An application proposing this work is current under review.

D. Publications Since January 1984

1. (*) Buchanan, B.G. and Shortliffe, E.H.: *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Addison-Wesley, Reading, MA., 1984. [book]

2. (*) Clancey, W.J. and Shortliffe, E.H.: *Readings in Medical Artificial Intelligence: The First Decade.* Addison-Wesley, Reading, MA., 1984. [book]

3. Clancey, W.J. and Shortliffe, E.H.: *Strategies for medical knowledge engineering: Lessons from the first decade.* To appear in the Proceedings of the AAMSI Congress 85, San Francisco, CA., May 1985.

4. Differding, J.C.: *The OPAL interface: General Overview.* Working paper. August 1984.

5. (*) Fagan, L.: *New Directions for Expert Systems: Examples from the ONCOCIN Project.* To appear in the Proceedings of AAMSI Congress 85, San Francisco, CA., May 1985.

6. (*) 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.* Submitted for publication, May 1985.

7. (*) Kahn, M.G., Ferguson, J., Shortliffe, E.H., Fagan, L.: *An approach for structuring temporal information in the ONCOCIN system.* To appear in the

Proceedings of the Symposium on Computer Applications in Medical Care, Baltimore, MD., November 1985.

8. (*) 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*. Submitted for publication, March 1985.

9. (*) Lane, C.D., Differding, J.C., Shortliffe, E.H.: *Design of a graphic interface for a medical expert system*. (Memo KSL-85-15). Working paper.

10. (*) Langlotz, C., Fagan, L., Tu, S., Williams, J., Sikic, B.: *ONYX: An architecture for planning in uncertain environments*. To appear in the Proceedings of International Joint Conference on Artificial Intelligence, Los Angeles, CA., August 1985.

11. (*) Langlotz, C.P. and Shortliffe, E.H.: *Adapting a consultation system to critique user plans*. In Developments in Expert Systems, (M. Coombs, ed.), pp. 77-94, London: Academic Press, 1984.

12. (*) Musen, M., Langlotz, C., Fagan, L., Shortliffe, E.H.: *Rationale for knowledge base redesign in a medical advice system*. To appear in the Proceedings of AAMSI Congress 85, San Francisco, CA., May 1985.

13. Shortliffe, E.H.: *The science of biomedical computing*.Medical Informatics, Vol.9, Nos. 3/4, 185-193 (1984).

14. (*) Shortliffe, E.H.:*Reasoning methods in medical consultation systems: artificial intelligence approaches* (tutorial). Computer Programs in Biomedicine 18:5-14 (1984).

15. Shortliffe, E. H.: *Explanation capabilities for medical consultation systems* (tutorial). Proceedings of AAMSI Congress 84 (D. Lindberg and M. Collen, Eds.), pp. 193-197, San Francisco, May 1984.

16. Shortliffe, E.H. and Fagan, L.M.: *Artificial intelligence: the expert systems approach to medical consultation*. Proceedings of the 6th Annual International Symposium on Computers in Critical Care and Pulmonary Medicine, Heidelberg, Germany, June 1984.

17. (*) Shortliffe, E.H.: *Update on ONCOCIN: A chemotherapy advisor for clinical oncology*. Proceedings of the Symposium on Computer Applications in Medical Care, November 1984.

18. (*) Tsuji, S. and Shortliffe, E.H.: *Graphics for knowledge engineers: a window on knowledge base management* (Memo KSL-85-11). Submitted for publication, April 1985.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

A great deal of interest in ONCOCIN has been shown by the medical, computer science, and lay communities. We are frequently asked to demonstrate the program to Stanford visitors (both the prototype system running in the clinic and the newer work transferring the system to professional workstations). We also demonstrated our developing workstation code in the Xerox exhibit in the trade show associated with AAAI-84 in Austin, Texas. Physicians have generally been enthusiastic about ONCOCIN's potential. The interest of the lay community is reflected in the frequent requests for magazine interviews and television coverage of the work. Articles about MYCIN and ONCOCIN have appeared in such diverse publications as *Time* and *Fortune*, whereas ONCOCIN has been featured on the "NBC Nightly News", the PBS

"Health Notes" series, and "The MacNeil-Lehrer Report." Due to the frequent requests for ONCOCIN demonstrations, we have produced a videotape about the ONCOCIN research which includes demonstrations of our the 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.

Our group also continues to oversee the MYCIN program (not an active research project since 1978) and the EMYCIN program. Both systems continue to be in demand as demonstrations of expert systems technology. MYCIN been demonstrated via networks at both national and international meetings 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 start out 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 on in July of 1983 in *Computers and Biomedical Research.*

B. Sharing and Interaction with Other SUMEX-AIM Projects

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

C. Critique of Resource Management

Our community of researchers has been extremely fortunate to work on a facility that has continued to maintain the high standards that we have praised in the past. The staff members are always helpful and friendly, and work as hard to please the SUMEX community as to please themselves. As a result, the computer is as accessible and easy to use as they can make it. More importantly, it is a reliable and convenient research tool. We extend special thanks to 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. *To transfer the oncology prototype from its current research computer to a professional workstation that provides a model for cost-effective dissemination of clinical consultation systems.* To meet this specific aim we will 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 coming year, we will:

   - Complete our OPAL protocol entry system
   - Continue entry of additional protocols, hopefully at the rate of one protocol/month (including testing)
   - Place a version of the OPAL protocol entry system into the clinic for use by physicians as a graphical reference guide to the protocols.

3. *To introduce ONCOCIN gradually for ongoing use so that by mid-1986 two professional workstations will be available in the oncology clinic to assist in the management of cancer patients.* During the next year, we will:

   - Implement the first workstation-based ONCOCIN system for use by physicians in the oncology clinic by the end of the calendar year 1985, adding a second workstation within a few months thereafter
   - Continue to operate the DEC-2020 version to maintain continuity of support in the clinic setting until the workstation version is fully operational.

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 is shifting to Xerox workstations, the SUMEX 2060 and the 2020 continue to be key elements in our research plan. The programs all make assumptions regarding the computing environment in which they operate, and the ONCOCIN prototype currently used in the clinic depends upon proximity to the DEC 2020 which enables us to use a 9600 baud interface.

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

The main development of our project will continue to take place on Dandelion lisp machines that we have purchased or have been donated by XEROX corporation. We also have special needs for more computing power for our ONYX therapy planning research, and have been able to share an upgraded Dandelion loaned by SUMEX for this work.

C. Requirements for Additional Computing Resources

The acquisition of the DEC 2020 by SUMEX was crucial to the growth of our research work. It has insured high quality demonstrations and has enabled us to develop a system (ONCOCIN) for real-world use in a clinical setting. As we have begun to develop systems that are potentially useful as stand-alone packages (i.e., an exportable

ONCOCIN), the addition of personal workstations has provided particularly valuable new resources. We have made a commitment to the smaller Interlisp-D machines (Dandelions) produced by Xerox, and our work will increasingly transfer to them over the next several years. Our 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 address space is permitting development of the large knowledge base that ONCOCIN requires. The graphics capability of the workstations has also enabled us to develop new methods for presenting material to naive users. In addition, the D-machines have provided a reliable, constant "load-average" machine for running experiments with physicians and doing development work. The development of ONCOCIN on the Dandelion will demonstrate the feasibility of running intelligent consultation systems on small, affordable machines in physicians' offices and other remote sites.

D. Recommendations for Future Community and Resource Development

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

Responses to Questions Regarding Resource Future

"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.

"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 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.

"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.

"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 research activity to LISP machines (currently Xerox 1108'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 [16].

B. Medical Relevance

The molecular structure of proteins is essential for understanding many problems of medicine at the molecular level, such as the mechanisms of drug action. Using NMR data from proteins in solution will speed up the determination.

C. Highlights of Progress

We have constructed a prototype of such a program, called PROTEAN, designed on the blackboard model [7], [12]. It is implemented in BB1 [13], a framework system for building blackboard systems that control their own problem-solving behavior [14](see discussion of BB1 above). We have coupled the reasoning program with an IRIS graphics terminal (shared with SUMEX) which displays protein structures at different levels of detail. This provides a visual understanding of how the program is behaving, which is essential for this problem.

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 partitions the problem of positioning all of a protein's constituent structures within a global coordinate system into sub-problems of positioning 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. It also

introduces additional constraints in the global characteristics of internally constrained partial solutions. For example, the conformations of partial protein solutions constrain their composability with other partial solutions.

2. The problem-solver attempts to solve sub-problems and coordinate solutions at multiple levels of abstraction, where lower levels of abstraction partition solution elements with finer granularity. For example, PROTEAN operates at three levels of abstraction. At the "Solid" level, it positions elements of the protein's secondary structure: alpha-helices, beta-sheets, and random coils. At the "Blob" level, it positions elements of the protein's primary structure of amino acids: peptide units and side-chains. At the "Atom" level, it positions the protein's individual atoms. Partial solutions at higher levels of abstraction reduce the combinatorics of search at lower levels. Conversely, tightly constrained partial solutions at lower levels introduce new constraints on higher-level solutions.

3. The problem-solver forbears hypothesizing specific partial solutions for a sub-problem in favor of preserving 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 only as constrained as the data are constraining. It also accommodates incompatible constraints by permitting disjunctive sub-families. For PROTEAN, disjunctive sub-volumes imply that the associated structure lies within any one of the sub-volumes or, if the structure is mobile, that it may move from one sub-volume to another.

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, successively restricting the spatial volumes within which they may lie. Independent application of different constraints finesses the problem of integrating qualitatively different kinds of constraints by simply integrating their results. In addition, successive restriction of the family of solutions obviates guessing which specific solutions within a family are likely to be consistent with subsequently applied constraints and the otherwise inevitable back-tracking.

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. Toleration of overlapping partial solutions is another accommodation of incomplete or incompatible constraints and potentially dynamic solutions. For PROTEAN, overlapping volumes for two protein 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 (see section 3.2.1). It also

provides a foundation for the problem-solver's explanation of problem-solving activities and intermediate partial solutions (see section 3.2.2) and for its learning of new control heuristics (see section 5.5).

The current version of PROTEAN has six knowledge sources that demonstrate the reasoning techniques described above. These knowledge sources develop partial solutions that position multiple helices at the Solid level and refine those helices at the Blob level. Proposed work will introduce knowledge sources that operate on other protein structures at the Solid level, as well as knowledge sources that apply the reasoning techniques at the Blob and Atom levels. We also will investigate emergent constraints entailed in reliable partial solutions, composition of partial solutions into complete solutions, and intelligent control.

D. Relevant Publications

1. Erman, L.D., Hayes-Roth, B., Lesser, V.R., Reddy, D.R.:*The HEARSAY-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty*. ACM Computing Surveys 12(2):213-254, June, 1980.

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

3. 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.

4. Hayes-Roth, B.:*A Blackboard Architecture for Control*. Artificial Intelligence In Press, 1985.

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

6. 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.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

Several members of Prof. Jardetzky's research group are involved in this research.

B. Interactions with other SUMEX-AIM projects

Robert Langridge was visiting at Stanford last year, and informal discussions with him and his group have continued in this year.

C. Critique of Resource Management

The SUMEX staff has continued to be most cooperative in getting this project started. Without their persistence, we would not have been able to obtain Ethernet software for the IRIS graphics terminal from Xerox.

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 six knowledge sources that demonstrate the reasoning techniques described above. These knowledge sources develop partial solutions that position multiple helices at the Solid level and refine those helices at the Blob level. The proposed research would expand PROTEAN to include knowledge sources that:

1. construct partial solutions combining helices, beta sheets, and random coils at the Solid level;
2. merge highly constrained partial solutions at the Solid level;
3. refine Solid level solutions in terms of the relative positions of constituent peptide units and side chains at the Blob level;
4. further restrict the relative locations of peptide units and side chains relative to one another at the Blob level;
5. propagate emergent constraints at the Blob level back up to the Solid level to further restrict the relative positions of superordinate helices, beta sheets, and random coils;
6. refine Blob level solutions at the Atom level;
7. further restrict the relative locations of atoms relative to one another;
8. propagate emergent constraints at the Atom level back up to the Blob level to further restrict the relative positions of superordinate peptide units and side chains.

The research will also develop a set of control knowledge sources to guide PROTEAN's application of constraints to identify the family of legal protein conformations as efficiently as possible. And we expect to improve the graphics interface to provide more functionality and options for viewing partial structures.

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 the Blackboard model and knowledge structures most flexibly and quickly.

C. Need for other computing resources

In this stage of development we need more computer cycles and hope to have access to additional D-machines. We expect to upgrade the Silicon Graphics IRIS terminal to a workstation for more efficiency in the subprograms doing computational geometry.

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

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

The fundamental objective of the ARAMIS Project 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.

B. Medical Relevance and Collaboration

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

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

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

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

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

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

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

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

Despite the expenditure of billions of dollars over recent years for randomized controlled trials (RCT's) designed to answer these and other questions, answers have been slow in coming. RCT's are expensive 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.

C. Highlights of Research Progress

C.1 April 1984 to April 1985

Our primary accomplishments in this period have been the following:

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

2) carrying out and publishing the study of the effect of prednisone on serum cholesterol on this expanded database,

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

4) completion of a System Programmer's Manuals and User's Manual

5) initiation of transfer of RADIX to Xerox 1108 personal work stations.

C.1.1 Modifications to RADIX for the enlarged database

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

C.1.2 Prednisone/chlosterol study on enlarged database

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

An extensive paper describing the RADIX System and reporting the results of the prednisone/cholesterol study has been submitted to a major medical journal for publication.

C.1.3 Publish description of 2-stage regression method

A detailed description of the 2-stage regression method used by us for the above study has been sent to a major statistical journal for publication.

C.1.4 Documentation

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

C.1.5 Transer of RADIX to D-Machines

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

C.1.6 Other accomplishments

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

In addition, new work on the theory of medical knowledge representation is described below.

C.2 Research in Progress

Our current work is focusing on problems involved in the representation of medical knowledge. Specifically, we are developing new methods for representing medical causal relationships. These have been represented in most other systems as simply binary relationships with conditional probabilities or certainty factors. In our project we are exploring the representation of causal relationships using categorical, rank, and real-valued relationships, as well as binary ones. We anticipate that these relationships will a) lend greater accuracy to predictions and diagnoses made by medical consultation systems, and b) will enable medical knowledge bases to be more compact and perspicuous.

In addition to this theoretical work, we are also pursuing two applications. First, we are developing a system for using a medical knowledge base to summarize a patient's time-oriented record. That is, our intended system will take as input a table of signs, symptoms, and lab values of the patient over time and will transform this into a time-oriented summary of arbitrary detail. This application draws upon our existing work in representation of causal relationships and in labeling time-oriented records.

Our second application involves the development of methods for automating the discovery of new relationships from time-oriented patient records. Here, we have elaborated a number of methods that we intend to exploit in a newly designed version of our discovery module. These methods take advantage of pre-existing medical knowledge by using analogical reasoning. We expect that this work will be facilitated by our recent acquisition of the KEE knowledge representation system, courtesy of Intellicorp, for use on our Xerox 1108's.

D. Publications

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

2. Blum, R.L.: *Prednisone Elevates Cholesterol: An Automated Study of Longitudinal Clinical Data.* (Submitted to the Annals of Internal Medicine.)

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

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

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

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

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

8. Blum, R.L. and Wiederhold, G.: *Studying hypotheses on a time-oriented database: An overview of the RX project.* Proc. Sixth SCAMC, IEEE, Washington D.C., October, 1982.

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

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

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

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

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

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

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

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

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

18. Kuhn, 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.

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

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

21. 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.

22. 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.

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

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

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

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

B. Interactions with Other SUMEX-AIM Projects

This past year, in moving our work to the Xerox 1108's, we have had frequent consultations with members of the Oncocin staff and have made use of several utility programs developed by them including hash file facilities and programs facilitating the tabular display of data.

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

C. Critique of Resource Management

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

The SUMEX resource management continues to be accessible and 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 --

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

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

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

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

B. Justification and Requirements for Continued Use of SUMEX

Computerized clinical data banks possess great potential as tools for assessing the efficacy of new diagnostic and therapeutic modalities, for monitoring the quality of health care delivery, and for support of basic medical research. Because of this potential, many clinical data banks have recently been developed throughout the United States. However, once the initial problems of data acquisition, storage, and retrieval have been dealt with, there remains a set of complex problems inherent in the task of accurately inferring medical knowledge from a collection of observations in patient records. These problems concern the complexity of disease and outcome definitions, the complexity of time relationships, potential biases in compared subsets, and missing and outlying data. The major problem of medical data banking is in the reliable inference of medical knowledge from primary observational data.

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

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

C. Recommendations for Resource Development

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

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 7/86, e.g., continue like it is, discontinue support of the central machine, act as a communications crossroads, develop software for user community workstations, etc.

A: In our opinion, the SUMEX 2060 should continue to be supported. The machine continues to be of value to us for text-editing (TV edit and emacs) and 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 forseeable 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 forseeable 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. CADUCEUS Project

CADUCEUS Project

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

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

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

A major achievement of this research undertaking has been the design of a program called INTERNIST-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-1 has performed at the level of the skilled internist, but the experience has high-lighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

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

The INTERNIST-1 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. In addition, selected other medical schools are employing the INTERNIST-1 knowledge base for medical student and house staff education.

----Accomplishments this past year

During 1983-84, under the supervision of Drs. Miller and Myers, Dr. Michael First, a former University of Pittsburgh medical student with extensive experience working in the Decision Systems Laboratory, developed a program called QUICK (QUick Index into Caduceus Knowledge), a prototypical electronic textbook of medicine utilizing the INTERNIST-1 knowledge base as its foundation. A paper describing QUICK, including an informal trial evaluating its utility, appears in the April 1985 issue of Computers and Biomedical Research. The residents in Internal Medicine who were given access to QUICK rated it favorably as a source of medical information. All three hospitals participating in the evaluation of QUICK have requested that they be given continued access to the program. An effort is being made to adapt QUICK to the IBM-PC for easier use by physicians.

From 1981 through 1983, Dr. Miller, under NLM New Investigator Award 5R23-LM03589, developed a clinical patient case simulator program, CPCS. The goal of the project was to build a program and knowledge base capable of constructing, de novo, logically consistent and clinically plausible artificial patient case summaries. Such a program would be useful in helping medical students to broaden their diagnostic skills. The program might also be used in generating cases for testing purposes, as this is now done manually by the National Board of Medical Examiners for their certification examinations. CPCS was a successful feasibility study; its performance has not yet been formally evaluated. Plans have been made to convert the entire INTERNIST-1 knowledge base into the format used by CPCS, and to add a better representation of time to the CPCS program and knowledge base.

Drs. Miller and Myers have developed, as part of the CPCS project, a new format for the internal medicine knowledge base. The specific details of this format have been described in previous progress reports. We have, in a period of three to four man-months, converted on paper the INTERNIST-1 knowledge base for liver diseases into the new format. This represents about one-sixth of the entire INTERNIST-1 knowledge base.

Dr. Miller has written an editor program to enter and maintain the new knowledge base, using Franz Lisp. At present, that editor program has been used to construct some 15-17 diagnoses from the INTERNIST-1 liver diseases. This includes creation of some 50-70 facets describing the underlying pathophysiology. A total of 200-300 findings have been entered into the new knowledge base, and because of their complexity, they correspond to 400-600 INTERNIST-1 style manifestations. During the past year, two fellows in Computer Medicine, Drs. Lynn Soffer and Fred Masarie, have converted all INTERNIST-1 findings into the new format required by CPCS.

Dr. Miller has also written, over the past year, a new diagnostic program which uses the information in the new knowledge base as a substrate for making diagnoses in internal medicine. The program's behavior is roughly comparable to that of INTERNIST-1 on similar cases in the limited problem domain currently available for testing. This remains an area of continued research activity.

In addition to the aforementioned work in internal medicine, Drs. Gordon Banks and John Vries have been working on the development of a neurological diagnostic component for CADUCEUS. Dr. Banks has developed a neuroanatomic database which contains spatial descriptors for nearly 1,000 neuroanatomic structures and contains information as to their blood supply and function. This database will allow anatomic localization of neurologic lesions. Some of this work for the peripheral nervous system has been done previously by students in our laboratory. The approach to the central nervous system has been to design a set of "symbolic coordinates". In constructing the neuroanatomic database, the human body, including the nervous system, is conceptually partitioned into a set of cubes (boxes). Attached to each cube LISP atom are lists of all of the anatomic structures that are completely and partially contained within the cube, as well as the blood supply to the region. This structure facilitates rapid retrieval of the location of a given anatomic structure as well as rapid localization of possible areas of involvement when there is evidence of dysfunction of one or more neural systems.

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

The neuroanatomic database has been successfully complemented on the VAX 11/780. Efforts are currently underway to implement the system on lower cost AI workstations such as the SUN and the PERQ.

Dr. Vries has continued to work on an image processing system based on "octree" encoding. Sean McLinden has developed an interface to the General Electric 9800 series CT scanner that permits direct input of data from the scanner to the octree system. The octree system output consists of 3 dimensional shaded images of CT objects at 1 mm resolution. Three dimensional images containing 2 million pixels can be scaled, translated, and rotated by the system in 30-60 seconds.

An interface to the neuroanatomic database has also been developed that maps the 27-ary tree representational scheme of the database into an octree representational scheme. This has been used to implement an interactive program that allows a user to generate a three dimensional image of the brain by logically ORing database objects.

A prototype system for the automated diagnosis of CT scans has also been implemented. The system uses the flavors package, and the RUP truth maintenance system to reason about the distribution of CT densities in quadtrees (2 dimensional representations) or octrees (3 dimensional representations). Such a system might ultimately provide CADUCEUS with direct access to the diagnostic information in neuro images.

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 and the pediatrics knowledge base has continued to grow.

----Research in progress

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

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

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

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

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

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

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

D. List of relevant publications

1. First, M.B., Soffer, L.J., Miller, R.A.:*QUICK (Quick Index to Caduceus Knowledge): Using the Internist-1/Caduceus knowledge base as an an electronic textbook of medicine*. Comput. Biomed. Res. April 1985.

2. Miller, R.A.: *Internist-1/CADUCEUS: Problems Facing Expert Consultant Programs. Methods of Information in Medicine*. Schattauer, Stuttgart New York, Vol. 23, No. 1, January 1984, pp. 9-14.

3. Miller, R.A.: *A Computer-based Patient Case Simulator. Clinical Research*. 1984, 32:651A. (abstract).

4. Miller, R.A., Schaffer, K.F., Meisel, A.: *Ethical and legal issues related to the use of computer programs in clinical medicine*. Annals of Internal Medicine. 1985, 102:529-536.

5. Myers, J.D.: *Educating future physicians: Something old, Something new*. Ohio State Univ. Proceedings of Symposium, Medical Education in the 21st Century. (in press.)

6. 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, Univ. of California Press (in press).

7. Pople, H.E.: *CADUCEUS: An Experimental Expert System for Medical Diagnosis. IN The AI Business*. Edited by Patrick H. Winston and Karen A. Prendergast. 1984, pp. 67-80.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A, B. Medical Collaborations and Program Dissemination Via SUMEX

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

C. Critique of Resource Management

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

III. RESEARCH PLANS

A. Project Goals and Plans

Continued effort to complete the medical knowledge 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 CADUCEUS on the VAX. This sequence retains the operative capability of INTERNIST-1 as a computerized "textbook of medicine" for educational purposes.

B. Justification and Requirements for Continued SUMEX Use

Our use of SUMEX will obviously decline with the installation of our VAX and the use of personal work stations. 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 CADUCEUS is developed here.

Our best prediction is that our project will require continued access to the 2060 for the next two to three years 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 VAX and personal work stations such as Symbolics. The imposition of fees for the use of SUMEX facilities would seem to involve unnecessary book-keeping and probably would detract from the use of SUMEX, which is currently so efficient and pleasant.

Our team hopes to remain as a component of the SUMEX community and to share experiences and developments.

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

Our predictable needs in this area will be met by our dedicated VAX computer and newly acquired personal work stations.

D. Recommendations for Future Community and Resource Development

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

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 three 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, in Press) and have completed several experiments evaluating their theory (Polson & Kieras, 1984, 1985).

B. Technical Goals

The CLIPR project consists of two subprojects. The first, the text comprehension project, is headed by Walter Kintsch and is a continuation of work on understanding of connected discourse that has been underway in Kintsch's laboratory for several years. The second, the device complexity project is headed by Peter Polson in collaboration with David Kieras of the University of 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, in Press)

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by 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.

C. Medical Relevance and Collaboration

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

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, 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.

Several other psychologists have either used or shown an interest in using an early version of the prose comprehension model, including Alan Lesgold of SUMEX's SCP project, who is exporting the system to the LRDC Vax. We have also worked with James Greeno -- another member of the SCP project -- on a project that will integrate this model with models of problem solving developed by Greeno and others at the University of California, Berkeley. Needless to say, all of this interaction has been greatly facilitated by the local and network-wide communication systems supported by SUMEX. The mail system, of course, has also enabled us to maintain professional contacts established at conferences and other meetings, and to share and discuss ideas with these contacts.

D. Progress Summary

The 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 third 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 (in Press). They have also completed several experiments evaluating the theory (Polson & Kieras, 1984, 1985) 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.

E. List of Relevant Publications

1. Fletcher, R. C.: *Understanding and solving word arithmetic problems: A computer simulation.* Technical Report N0. 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, In Press.

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

4. Kintsch, W. and Greeno, J.G.:*Understanding and solving word arithmetic problems.* Psychological Review, 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.* Proceedings of the CHI 1985 Conference on Human Factors in Computing. San Francisco, April 1985.

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.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Sharing and Interactions with Other SUMEX-AIM Projects

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

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

B. Critique of Resource Management

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

III. RESEARCH PLANS

A. Long Range Projects Goals and Plans

The 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. However, we consider our continued access to SUMEX critical for the successful continuation of these projects.

Access to SUMEX provides us with continued contact with the SUMEX community, which is especially critical for the prose comprehension project. Knowledge representation languages, e.g. UNITS, and other tools developed by SUMEX are critical for this project. Alternative sources of such software are typically unsatisfactory because the systems have only been developed for use on one project and are typically very poorly documented and less than completely debugged. We hope that our continued membership in the community will be offset by the input that we have been and will continue to provide to various projects: our relationship has been symbiotic, and we look forward to its continuation.

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 four computing systems for the VAX 11/780, and three Xerox 1108s, one of which is at the University of Michigan. The VAX is used primarily 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 two critical needs. The first is communication, which we discussed in the preceding paragraph. The second is technical advice and access to various knowledge representation languages like UNITS.

We envisage our future needs to be communication currently served by the SUMEX 2060 and technical advice and necessary software provided by the SUMEX staff.

D. Recommendations for Future Community and Resource Development

Our future needs are for the SUMEX-AIM resource to act as a communications crossroad and to develop software and provide technical support for user community work stations. We have no preferences as to how such services are provided either with a communication server on the network or with the central machine like the current 2060.

We will continue to need access to the SUMEX-AIM 2060 in order to access communication networks and to interact with the SUMEX-AIM staff and community.

If communications and access to the staff are provided through some other mechanism, then we would no longer need access to the 2060.

We would be willing to pay fees for using SUMEX communication resources if required by NIH. However, our willingness is price sensitive. Any charges over $1,000 a year would mean we should communicate with people directly by long-distance telephone.

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 intended to be a collaboration between Dr. Blaschke at Stanford and Dr. Speedie at the University of Maryland. Dr. Speedie provides the expertise in the area of artificial intelligence programming. Dr. Blaschke provides the medical expertise. 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 has focused on exploration of the problem of designing the MENTOR system. Work has begun on constructing a system for monitoring potassium in patients with drug therapy that can adversely affect potassium. Antibiotics, dosing in the presence of renal failure, and digoxin dosing have been identified as additional topics of interest.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

This project represents a collaboration between faculty at Stanford University Medical Center and the University of Maryland School of Pharmacy in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, facilitates this collaboration of geographically separated project participants by allowing development work on a central machine resource 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 issues related to the formulation of the previously mentioned proposal. We expect interactions with other projects to increase significantly once the groundwork has been laid and issues directly related to AI are being addressed. Given the geographic separation of the investigators, the ability to exchange mail and programs via the SUMEX system as well as communicate with other SUMEX-AIM projects is vital to the success of the project.

C. Critique of Resource Management

To date, the resources of SUMEX have been fully adequate for the needs of this project. The staff have been most helpful with any problems we have had and we are quite satisfied with the current resource management. The only concerns we have relate to the state of the documentation on the system and the response time while using TYMNET from the Baltimore, Maryland area. While most aspects of the system are documented the path to a specific piece of information can be somewhat longer than one might expect. With respect to TYMNET, there are often up to 7 second pauses in the middle of transmissions. This can become quite annoying when trying to work with anything more than small bodies of text.

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.

1985 will be spent on prototype development in four content areas, design and implementation of the basic knowledge representation and reasoning mechanisms and preliminary 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.

The resources of SUMEX are central to the execution of the MENTOR project. A major component of the proposal was access to SUMEX resources and without it, the chances of funding would have been much less. 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 intend to transport the prototype system to a dedicated hardware system that can fully support the the planned system and which can be integrated into the SUMC Hospital Information System. However, no firm decisions have been made about the requirements for this system since many specification and design decisions remain to be made.

D. Recommendations for Future Community and Resource Development

In the brief time we have been associated with SUMEX, we have been generally pleased with the facilities and services. However, it is 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 nation-wide 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 that Stanford. Specifically, the problem of slow throughput on TYMNET needs to be addressed for those users who do not have authorized access to ARPANET. 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 to be replaced by an efficient, high-speed communications server.

IV.B.4. Rutgers Research Resource

Rutgers Research Resource--Artificial Intelligence in Medicine

Principal Investigators:
Casimir Kulikowski, Sholom Weiss
Rutgers University, New Brunswick, New Jersey

I. SUMMARY OF RESEARCH PROGRAM

A. Goals and Approach

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

An essential part of the resource is directed to methodological problems of knowledge representation and to the development of computer-based systems for acquiring, managing, and improving knowledge bases, and for constructing expert reasoning models in medicine. Equally fundamental are the problems of how to best use knowledge bases and models in processes of interpretation/diagnosis, planning, theory formation, simulation, and effective man-machine communication. These are problems we are studying in the Resource in the context of several system building efforts that address themselves to specific tasks of clinical decision-making and model development and testing.

Resource activities include research projects (collaborative research and core research) training/dissemination projects, and computing services in support of user projects.

B. Medical Relevance and Collaborations

In 1984-85 we continued the development of several versatile systems for building and testing consultation models in biomedicine. The EXPERT system has had many of its capabilities enhanced in the course of collaborative research in the areas of rheumatology, ophthalmology, and clinical pathology.

In *ophthalmology* we have developed a knowledge representation scheme for treatment planning which is both natural and efficient for encoding the strategies for choosing among competing and cooperating treatment plans. This involves a ranking of treatments according to their characteristics and desired effects as well as contraindications. A diagnosis and treatment planning program for ocular herpes was developed using this scheme. Our main collaboration continues to be with Dr. Chandler Dawson of the Proctor Foundation, UCSF.

In *rheumatology*, the model for rheumatological diseases now includes detailed diagnostic criteria for 26 major diseases. The management advice and treatment planning has been developed further. The Resource researchers have developed new representational elements for EXPERT in response to the needs of the rheumatology research Politakis originally developed a coordinated system called SEEK (System for

Empirical Experimentation with Expert Knowledge) which provides interactive assistance to the human expert in testing, refining and updating a knowledge base against a data base of trial cases. A generalized version of SEEK, SEEK2, has been developed during the past year. Dr. Lindberg of the National Library of Medicine, and Dr. Sharp, of the University of Missouri are the project leaders in developing the rheumatology knowledge base for this effort.

In *clinical pathology* our main collaboration has been with Dr. Robert Galen (Cleveland Clinic Foundation), with whom we have developed the serum protein electrophoresis model which is incorporated into an instrument a scanning densitometer. This instrument with interpretive reporting capabilities has now been on the market for over a year, is located at several hundred clinical sites. We are making good progress developing a knowledge based system for the interpretation of CPK/LDH isoenzymes.

In biomedical modeling applications we are experimenting with several prototype models for giving advice on the interpretation of experimental results in the field of enzyme kinetics, in conjunction with Dr. David Garfinkel. His PENNZYME program has been linked to a model in EXPERT, which allows the user to interpret the progress of the model analysis, and a framework for the design of experiments in this domain has been formulated.

C. Highlights of Research Progress

Research has continued on problems of representation, inference and control in expert systems. Emphasis has been placed this year on problems of knowledge base acquisition, empirical testing and refinement of reasoning (the SEEK2 system). From a technological point of view the market availability of the interpretive reporting version of a scanning densitometer, and the development of models for eye care consultation that run on microprocessor systems (Apple IIe, IBM-PC) represents an important achievement for AIM research in showing its practical impact in medical applications. This was recognized by the award of a scientific exhibit prize at the Academy of Ophthalmology Annual Meeting in November 1983.

- Knowledge Base Refinement: SEEK is a system which has been developed to give interactive advice about rule refinement during the design of an expert system. The advice takes the form of suggestions for possible experiments in generalizing and specializing rules in an expert model that has been specified based on reasoning rules cited by a human expert. Case experience, in the form of stored cases with known conclusions, is used to interactively guide the expert in refining the rules of a model. The design framework of SEEK consists of a tabular model for expressing expert-modeled rules and a general consultation system for applying a model to specific cases. This approach has proven particularly valuable in assisting the expert in domains where the logic for discriminating two diagnoses is difficult to specify; and we have benefited primarily from experience in building the consultation system in rheumatology. During the past year a newer SEEK2 system a been developed that has enhanced capabilities including a more generalized knowledge base and an automatic pilot capability to proceed with knowledge base refinements.

- Technology Transfer: Important technology transfer milestones have also been achieved this year: the instrument interpretation EXPERT program for serum protein has been widely disseminated as has the Ocular Herpes Treatment Program.

D. Up-to-Date List of Publications

The following is an update of publications in the Rutgers Resource for the period 1983 and 1984 (only publications not listed in previous SUMEX annual reports are presented here).

1. Apte, C. and Weiss, S.: *An Approach to Expert Control of Interactive Software Systems, IEEE Transactions on Pattern Analysis and Machine Intelligence* in press (1985).

2. Ginsberg, A., Weiss, S., and Politakis, P.: *SEEK2: A Generalized Approach to Automatic Knowledge Base Refinement* to appear in the Proceedings of the 1985 International Joint Conference on Artificial Intelligence.

3. Weiss, S.M. and Kulikowski, C.A.: *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

4. Kastner, J., Weiss, S., Kulikowski, C., and Dawson, C.: *Therapy Selection in an Expert Medical Consultation System for Ocular Herpes Simplex* Computers in Biology and Medicine, Vol. 14, No. 3, pp. 285-301 (1984).

5. Dawson, C., Kastner, J., Weiss, S., Kulikowski, C.: *A Computer-based Method to Provide Subspecialist Expertise on the Management of Herpes Simplex Infections of the Eye*, Proceedings International Symposium On Herpetic Eye Diseases, Belgium (1984).

6. Galen, R. and Weiss, S.: *Predictive Value Calculator*, American Society of Clinical Pathologists, Clinical Chemistry # CC 84-4 (1984).

7. Kastner, J., Dawson, C., Weiss, S., Kern, K., Kulikowski, C.: *An Expert Consultation System for Frontline Health Workers in Primary Eye Care*, Journal of Medical Systems, Vol. 8, No. 5 (1984).

8. Kulikowski, C.A.: contributor to the Knowledge Acquisition chapter edited by B. Buchanan in the book *Building Expert Systems* (F. Hayes-Roth, et al., eds) Addison-Wesley, 1983.

9. Yao, Y. and Kulikowski, C.A.: *Multiple Strategies of Reasoning for Expert Systems*, Proc. Sixteenth Hawaii International Conference on Systems Sciences, pp. 510-514 , 1983.*

10. Kulikowski, C.A.: *Progress in Expert AI Medical Consultation Systems: 1980 - 1983* , Proc. MEDINFO '83 , pp. 499-502, Amsterdam, August 1983.*

11. Kastner, J.K., Weiss, S.M., and Kulikowski, C.A.: *An Efficient Scheme for Time-Dependent Consultation Systems*, Proc. MEDINFO '83, pp.619-622, 1983.*

12. Kulikowski, C.A.: *Expert Medical Consultation Systems*, Journal of Medical Systems, v.7, pp. 229-234, 1983.*

13. Weiss, S.M., Kulikowski, C.A., and Galen, R.S.: *Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program*, Journal of Clinical Laboratory Automation, v.3, pp. 383-387, 1983.*

14. Kastner, J.K., Weiss, S.M., and Kulikowski, C.A.: *An Expert System for Front-line Health Workers in Primary Eye Care*, Proc. Seventeenth Hawaii International Conference on Systems Sciences, pp. 162-166, 1984.*

15. Kulikowski, C.A.: *Knowledge Acquisition and Learning in EXPERT*, Proc. 1983 Workshop on Machine Learning, Univ. of Illinois,Champaign-Urbana 1983.

Indicate by an asterisk (*) that the resource was given credit.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Dissemination

The SUMEX-AIM facility provides a backup node where some of our medical collaborators can access programs developed at Rutgers. The bulk of the medical collaborative work outlined in I.B. above is centered at the Rutgers facility (the Rutgers-AIM node).

Dissemination activities continue to be an important responsibility of the Rutgers Resource within the AIM community. The following activities took place in the last year:

1. Tenth AIM Workshop (1983):

   Organized by Dr. Chandrasekaran, it was held at Ohio State University. It consisted of a series of presentations on AIM research and related work by members of the AIM community.

2. 1984 Hawaii International Conference On Systems Sciences:

   Dr. Weiss presented a paper on the expert system for front-line health workers, and Dr. Kulikowski chaired a session on knowledge based medical systems.

B. National AIM Projects at Rutgers

The national AIM projects, approved by the AIM Executive Committee, that are associated with the Rutgers-AIM node are the following:

1. INTERNIST/CADUCEUS project, headed by Dr. Myers and Dr. Pople from the University of Pittsburgh, has been using the Rutgers Resource as a backup system for development and experimentation.

2. Medical Knowledge Representation project, headed by Dr. Chandrasekaran from Ohio State University, is doing most of its research on the Rutgers system.

3. PURSUIT project, directed by Dr. Greenes from Harvard University, is doing most of its research on a Goal-Directed Model of Clinical Decision-Making at Rutgers.

4. Biomedical Modeling, by Dr. Garfinkel from the University of Pennsylvania.

5. Attending Project, directed by Dr. Perry Miller of the Yale Medical Center, is doing much of the research on critiquing a physician's plan of management at Rutgers.

6. MEDSIM project: This is a pilot project designed to provide resource-sharing and community building facilities for about 25 researchers in bio-mathematical modeling and simulation.

C. Critique of SUMEX-AIM Resource Management

Rutgers is currently using the SUMEX DEC-20 system primarily for communication with other researchers in the AIM community and with SUMEX staff, and also for backup computing in demonstrations, conferences and site visits. Our usage is currently running at less than 50 connect hours per year at SUMEX, with an overall connect/CPU ratio of about 30.

Rutgers is beginning to place more emphasis on the use of personal computers, and on network support needed to make these effective. SUMEX has been of significant help in their developmental efforts in networking workstation software.

III. RESEARCH PLANS

A. Project Goals and Plans

We are planning to continue along the main lines of research that we have established in the Resource to date. Our medical collaborations will continue with emphasis on development of expert consultation systems in rheumatology, ophthalmology and clinical pathology. The basic AI issues of representation, inference and planning will continue to receive attention. Our core work will continue with emphasis on further development of the EXPERT framework and also on AI studies in representations and problems of knowledge and expertise acquisition. We propose to work on a number of technology transfer experiments to micro processing that will be affordable by our biomedical research and clinical collaborators. We also plan to continue our participation in AIM dissemination and training activities as well as our contribution -- via the Rutgers computers -- to the shared computing facilities of the national AIM network.

B. Justification and Requirements for Continued SUMEX Use

Continued access to SUMEX is needed for:

1. Backup for demos, etc.

2. Programs developed to serve the National AIM Community should be runnable on both facilities.

3. There should be joint development activities between the staffs at Rutgers and SUMEX in order to ensure portability, share the load, and provide a wider variety of inputs for developments.

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

Our computing needs are based on a centralized computing resource accessible to distant users, and local workstations. We will continue to use Sumex for backup purposes.

D. Recommendations for Future Community and Resource Development

Use of personal computers and workstations is continuing to grow in the AIM community. We find that the biggest challenge is supporting these systems. Although some central computing will continue to be needed for communication and coordination, we believe that over the next few years all AIM research projects and even individual collaborators will come to have their own hardware. However many of these community members (particularly the collaborators) will not be in a position to support hardware or software on their own. We would certainly expect SUMEX to continue to provide expert advice in this area. However we believe it would be helpful for SUMEX to have a formal program to support smaller computers in the field. We envision this as including at least the following items:

- A central source of information on hardware and software that is likely to be of interest to the AIM community. SUMEX might want to become a distribution point for certain of this software, and even help coordinate quantity purchase of hardware if this proves useful.

- Assistance in support of hardware and software in the field. Depending upon the hardware involved, this might involve advice over the telephone or actual board-swapping by mail.

IV.B.5. SECS: Simulation & Evaluation of Chemical Synthesis

SECS - Simulation and Evaluation of Chemical Synthesis Project

Principal Investigator: W. Todd Wipke
Board of Studies in Chemistry
University of California
Santa Cruz, CA. 95064

Coworkers:

I. Kim	(Grad student)
M. Hahn	(Grad Student)
M. Yanaka	(Postdoctoral)
I. Iwataki	(Postdoctoral)
T. Okada	(Postdoctoral)

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

With the SECS project our long range goal is to develop the logical principles of molecular construction and to use these in developing practical computer programs to assist investigators in designing stereospecific syntheses of complex bio-organic molecules. Our second area of research, the XENO project, is aimed at improving methods for predicting potential biological activity of metabolites and plausibility of incorporation and excretion of metabolites.

B. Medical Relevance and Collaboration

The development of new drugs and the study of drug structure biological activity relationships depends upon the chemist's ability to synthesize new molecules as well as his ability to modify existing structures, e.g., incorporating isotopic labels or other substituents into bio-molecular substrates. The Simulation and Evaluation of Chemical Synthesis (SECS) project aims at assisting the synthetic chemist in designing stereospecific syntheses of biologically important molecules. The advantages of this computer approach over normal manual approaches are many: 1) greater speed in designing a synthesis; 2) freedom from bias of past experience and past solutions; 3) thorough consideration of all possible syntheses using a more extensive library of chemical reactions than any individual person can remember; 4) greater capability of the computer to deal with the many structures which result; and 5) capability of computer to see molecules in a graph theoretical sense, free from the bias of 2-D projection.

The objective of using XENO in metabolism studies is to predict the plausible metabolites of a given xenobiotic in order that they may be analyzed for possible carcinogenicity. Metabolism research may also find this useful in the identification of metabolites in that it suggests what to look for. Finally, one may envision applications of this technology in problem domains where one wishes to alter molecules in order to inhibit certain types of metabolism.

C. Highlights of Research Progress

C.1 SECS Project Developments

The focus of our work this year on SUMEX was conversion of our research programs from the SUMEX DEC-20 to a VAX 11/750 located in our research group. We were restricted to a 3% maximum cpu utilization on SUMEX which effectively precluded significant production work on SUMEX. We completed moving all files from SUMEX to our VAX 11/750 31 March 1985.

C.1.a SECS on VAX

The majority of the SECS program has been converted to the VAX in Fortran and is operational. A graphic driver for the Evans & Sutherland PS330 display system has been also added. New chemical transforms in heterocyclic chemistry have been written and debugged. Through our collaborations in Japan 1000 new chemical transforms have been added using an automatic ALCHEM transform writing program. All new developments in the SECS program will occur on the VAX version.

C.2 XENO Program Developments

The metabolic fate of various compounds in the human body is extremely complex, yet extremely important for it is known that through metabolism certain otherwise harmless compounds are converted into toxic and possibly carcinogenic agents. Because of this complexity it is difficult, looking at a given compound, to forecast potential biological activity of that given compound. The objective of this proposal is to develop a practical computer program by which a biochemist or metabolism expert can explore the metabolites of a given compound and be alerted to the plausible biological activity of each metabolite.

C.2.a Evaluation Study

We participated in an evaluation of XENO predictions of metabolism on four pre-manufacturing notice compounds from the U.S. Environmental Protection Agency, Office of Toxic Substances, in comparison with two panels of metabolism experts. These four compounds were selected from a list of six compounds considered by the EPA to be representative of the types and diversity of compounds they must evaluate. The limit of 3% cpu maximum utilization precluded evaluating more compounds.

The predictions of XENO were submitted to a third party as were the predictions from the two other panels of experts. The results from all three groups were then distributed and discussed at a meeting in Washington, DC.

In processing these four examples, the XENO program performed without crashing or errors. The graphical display equipment broke down during the study, but because XENO also permits teletype graphics, we were still able to complete the study. The total computer time used was approximately 15 minutes on a DEC 2060 system which is very little time for such analyses.

The results of the evaluation proved very interesting. The knowledge base of XENO was shown to be missing a couple transforms having to do with cleavage of C-S bonds, disulfide formation, and phosphorylation. These transforms have now been added to XENO, which required about 15 minutes, illustrating the simplicity of augmenting the knowledge base. But beyond the couple missing transforms, XENO correctly included all predictions by the experts and further suggested additional metabolites that might be present which the experts had not included.

XENO agreed with the experts more than the experts agreed with each other. The experts tended to approach the problems very narrowly, with just a few selected pathways. XENO tended to include the results of all the experts, approaching the problem more broadly. If the objective is risk assessment, the latter strategy is preferable. XENO also suggested some reasonable pathways, such as azo reduction in aryl-alkyl azo compounds, but the experts, having never seen results from such a compound concluded that because it had not been reported, it didn't occur. Now however an experimental study of azo reductase has been launched to determine what does happen with aryl-alkyl azo compounds.

Finally, as might be expected, the experts were biased against the computer expert system, and had greater difficulty seeing its potential than others involved in the risk assessment process.

C.2.b Molecular Model Builder

Over the past year we have begun a new project, to replace the molecular model builder in XENO with a faster and more general one. This will allow steric evaluation to be done more quickly and accurately. The goal of our project is to build a knowledge based program which can quickly and accurately create three-dimensional molecular models of organic molecules. Unlike other numerically oriented modelling programs, our program utilizes a large body of existing conformational data to infer preferred geometries. This knowledge base is the Cambridge Crystal file, which contains x-ray determined geometries for over 20,000 organic compounds.

The design work for the program was completed during the past year and now we are at the early stages of implementation. The program consists of the following individual modules:

1. A graphical front end facilitates input into the program and the display of results. The graphics package is a flexible visual tool for the chemist and runs on an Evans and Sutherland PS300 linked asynchronously to our VAX 750. It allows the easy construction and manipulation of both two-dimensional and crude three-dimension structures.

2. A perception module perceives the input structure for atom types, bond types, stereochemistry, bonding configuration, rings, and ring assemblies.

3. A search strategy generation module uses the perception data to formulate hierarchical rules, constraints, and goals used in searching the data base for possible structural knowledge to be used in model construction. Generation of the search strategy can be interactively guided by constraints and priorities defined by the user.

4. A construction module applies the knowledge found using a set of attachment rules and attempts to construct models which meet the initial constraints.

5. An evaluation module evaluates the models generated to determine the confidence level for the three-dimensional accuracy of each part of the model. This evaluation is based on criteria such as degree of analogy between previous precedent and current model.

Currently, the first, second and fourth modules of the program have been implemented and are being tested.

C.2.c Collaborative Efforts.

The co-operation between the groups at the University of Lund and the University of California Santa Cruz continues to prove fruitful for both parties. The SECS program, which was implemented in Sweden by Dr. Robert E. Carter after his visit to Santa Cruz in 1982, is still being used by both graduate and undergraduate students. Currently, SECS is hosted on a PDP-10 which is located 200 miles to the north of Lund. However, Lund is going to lose access to this machine in the foreseeable future. Fortunately, both Lund and Santa Cruz have purchased VAX machines, and Prof. Wipke has indicated that Lund will receive a VAX version of SECS in the near future.

Further cooperation was accomplished this winter when Dr. Dolata, formerly of Santa Cruz, and now at Lund, visited Prof. Wipke. Since Lund had obtained its VAX about 6 months previous to Santa Cruz, they had had time to build a repertory of useful programs and procedures. These were installed on the Santa Cruz VAX, thus improving the programming environment substantially.

In addition, Dr. Dolata gave a seminar on the current work in conformational analysis by symbolic reasoning which is under investigation at Lund, and received many thoughtful and helpful insights. A copy of the WIZARD conformational analysis system was provided for examination by the Santa Cruz group. Additionally, several papers to be published by Wipke and Dolata were discussed, and work was started on these papers.

Finally, with the upcoming installation of UUCP net on Lunds Vax, communication between UCSC and LU should be facilitated, so that even closer cooperation can be achieved.

The SECS project continues to have collaborations with the pharmaceutical industry which is adding chemical transforms and doing some joint program development, for example, Dr. Yanaka continued work started at Santa Cruz after he returned to Kureha Chemical in Japan and a paper has been prepared on that work.

In addition to collaboration with the SECS project, Dr. David Rogers at the University of Michigan writes: The SUMEX-AIM site has been a useful and necessary link for our AI research group at the University of Michigan to the ARPAnet community. Our work is an attempt to build a working system based on emergent structure appearing as the result of the statistical interaction of low-level subcognitive units; our work is being done on a network of SUN microcomputers using Franz Lisp. We appreciate the existence of SUMEX-AIM as an assist at keeping abreast with work at Stanford and other ARPA sites.

D. List of Current Project Publications

1. Wipke, W.T., and Rogers, D.: *Artificial Intelligence in Organic Synthesis. SST: Starting Material Selection Strategies. An Application of Superstructure Search.* J. Chem. Inf. Comput. Sci., 24:1 71-81, 1984.

2. Wipke, W.T., and Rogers, D.: *Rapid Subgraph Search Using Parallelism* J. Chem. Inf. Comput. Sci., 24:4 255-262 (1984).

3. Wipke, W.T.: "An Integrated System for Drug Design" in *The Aster Guide to Computer Applications in the Pharmaceutical Industry* Aster Publishing Co., Springfield, Oregon, 1984, pp 149-166.

4. Wipke, W.T.: *Computer Modeling in Research and Development*, Cosmetics and Toiletries, 99:Oct 73-82 (1984).

5. Wipke, W.T.: *Computer-Assisted Design of Organic Synthesis. ALCHEM: A Language for Representing Chemical Knowledge*, J. Chem. Info. Comput. Sci., 24, 0000 (1985).

6. Johnson, C.K., Thiessen, W.E., Burnett, M.N., Condran, P. Ronlan, A., Yanaka, M. and Wipke, W.T.: *Systematic derivation of chemical procedures for transforming surplus hazardous chemicals to useful products, J. of Hazardous Materials.* (In press, the appearance of this article has been delayed by Oak Ridge.)

7. Dolata, D.P.: *QED: Automated Inference in Planning Organic Synthesis* (Ph.D. dissertation). University of California, Santa Cruz, 1984.

8. Rogers, D.: *Artificial Intelligence in Organic Chemistry. SST: Starting Material Selection Strategies* (Ph. D. dissertation). University of California, Santa Cruz, 1984.

F. Research Environment

At the University of California, Santa Cruz, we have been previously connected to the SUMEX-AIM resource by a 4800 baud multiplexed leased line. Now we have disconnected that line and are using a VAX 11/750 as our host computer running the VMS operating system. We have a PS300 black/white vector graphic display which is driven by a serial line to the VAX. The SECS laboratory is located in 125 Thimann Laboratories, adjacent to the synthetic organic laboratories at Santa Cruz.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

SECS had been available in the GUEST area of SUMEX for casual users. SECS and XENO are no longer available through SUMEX. Access now must be through UCSC or by installation on the user's own computer.

Communication between SECS collaborators is facilitated by using SUMEX message drops, especially when time differences between the U.S. and Europe and Australia makes normal telephone communication difficult.

B. Examples of Cross-fertilization with other SUMEX-AIM Projects

The AILIST bulletin board has been used extensively for interacting with many projects and locating references for further information related to program design and AI technology. There are no longer any other chemical or biochemical projects on SUMEX so our interaction with the community is limited to AI technology interchange.

C. Critique of Resource Services

SUMEX-AIM gives us at UCSC, a small university, the advantages of a larger group of colleagues, and interaction with scientists all over the country. Since 1 April 1984, the computer response time has been very poor for the SECS project because our project was put in a separate class with a 3% cpu limitation. This was a very severe restriction which prevented short usage peaks from being averaged with other users. Projects in their final year should not be so restricted.

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

SECS 2.9 has been installed on the CompuServe computer networks for the past four years so anyone can access it without having to convert code for their machine. This has proved very useful as a method of getting people to experiment with this new technology. SECS also resides on the Medicindat machine at the University of Gothenborg, Sweden, and is available all over Sweden by phone. Similarly in Australia, SECS resides at the University of Western Australia and is available throughout Australia over CSIRONET. SECS has been installed at two locations in Japan. FSECS has been installed on a DEC-10 at Oak Ridge National Laboratory and serves for collaborative development of that approach with Carroll Johnson. PRXBLD has been disseminated to over 60 sites on various types of computers including DEC-10, DEC-20, IBM, VAX, PRIME, FUJITSU and Honeywell.

III. RESEARCH PLANS (4/85-4/86)

A. Near-Term Project Goals and Plans

Our planned use of the SUMEX resource is simply for message communication with collaborators. We will continue developing the SECS and XENO projects on the VAX 11/750 and incorporate graphics with the Evans and Sutherland PS300 system. A proposal is pending to add color displays to this system.

B. Justification and Requirements for Continued Use of SUMEX

We request to have continued access to SUMEX for receiving and sending messages to collaborators and for access to the important bulletin boards maintained on SUMEX. We may also need to retrieve some of our files archived on SUMEX since in moving ten years of research work off SUMEX it is possible we missed some key file which we will not recognize until we need it.

C. Needs Beyond SUMEX-AIM

In addition to our VAX, we are exploring graphic workstations to achieve a distributed environment since the VAX alone loads down very quickly. And we are seeking to add color to our Evans and Sutherland PS300.

D. Recommendations for Community and Resource Development

An important part of medicine is treatment of diseases with drugs. Drugs are chemicals--chemicals that were designed and synthesized by chemists. Since the termination of the DENDRAL project, there seems to be declining support for artificial intelligence applications in chemistry. We feel that support of this area is essential to the advancement of medicine in this country. The lack of chemists on NIH Research Resources computing peer review is contributing to the problem. Application of artificial intelligence in synthesis planning is one of the more successful current applications and it is now a high priority research area in many foreign countries. To maintain our lead in this technology, further funding is required.

Responses to Questions Regarding Resource Future

The SECS group feels that SUMEX should remain a communications center, but there is little need for it to attempt to grow the mainframe in an effort to supply cpu cycles to individual projects. It is now financially feasible for each project to have its own computer. But there is still a need for network access, knowledge sharing, file transfer, etc. SUMEX could serve this networking aspect with considerably less hardware and staff than it now has.

Since SUMEX no longer purports to serve a national community except for communication, there is no justification for continuing to grow the mainframe.

It is hard to see justification for SUMEX to develop workstation software since that is already being done commercially and since a similar proposal to RR to do same from San Diego was disapproved on the basis of it being inappropriate.

We expect to need access to SUMEX for message purpose only. That access is desired for probably two years or more, or until the UC network is operational. Currently much of the UC network is UNIX and VMS people can't currently connect.

Regarding the imposition of fees for service, I think that would be sad. There are already many networks that operate on a fee for service basis, i.e., Source, CompuServe, etc. If SUMEX had to be on a fee for service basis, it is unclear why the service might not better be handled by existing commercial vendors that have customer relations staff. It is unclear also that NIH grants would allow expenses for communication rather than hard computing.

Finally, just a note that I have appreciated the service of SUMEX, and the staff of SUMEX, although I did not appreciate the 3% limit under which we had to work last year.

IV.B.6. SOLVER Project

SOLVER: Problem Solving Expertise

Dr. P. E. Johnson
Center for Research in Human Learning
University of Minnesota

Dr. W. B. Thompson
Department of Computer Science
University of Minnesota

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

This project focuses upon the development of strategies for discovering and documenting the knowledge and skill of expert problem solvers. In the last several years, considerable progress has been made in synthesizing the expertise required for solving extremely complex problems. Computer programs exist with competency comparable to human experts in diverse areas ranging from the analysis of mass spectrograms and nuclear magnetic resonance (Dendral) to the diagnosis of certain infectious diseases (Mycin).

Design of an expert system for a particular task domain usually involves the interaction of two distinct groups of individuals, "knowledge engineers," who are primarily concerned with the specification and implementation of formal problem solving techniques, and "experts" (in the relevant problem area) who provide factual and heuristic information of use for the problem solving task under consideration. Typically the knowledge engineer consults with one or more experts and decides on a particular representational structure and inference strategy. Next, "units" of factual information are specified. That is, properties of the problem domain are decomposed into a set of manageable elements suitable for processing by the inference operations. Once this organization has been established, major efforts are required to refine representations and acquire factual knowledge organized in an appropriate form. Substantial research problems exist in developing more effective representations, improving the inference process, and in finding better means of acquiring information from either experts or the problem area itself.

Programs currently exist for empirical investigation of some of these questions for a particular problem domain (e.g. AGE, UNITS, RLL). These tools allow the investigation of alternate organizations, inference strategies, and rule bases in an efficient manner. What is still lacking, however, is a theoretical framework capable of reducing dependence on the expert's intuition or on near exhaustive testing of possible organizations. Despite their successes, there seems to be a consensus that expert systems could be better than they are. Most expert systems embody only the limited amount of expertise that individuals are able to report in a particular, constrained language (e.g. production rules). If current systems are approximately as good as human experts, given that they represent only a portion of what individual human experts know, then improvement in the "knowledge capturing" process should lead to systems with considerably better performance.

In order to obtain a broad view of the nature of human expertise, the SOLVER project includes studies in a variety of complex problem solving domains in addition to medicine. These include law, auditing, business management, plant pathology, and expert system design. We have observed that despite the apparent dissimilarities in these problem solving areas there is reason to believe that there are underlying principles of expertise which apply broadly. Our project seeks to investigate these principles and to create tools to make use of that knowledge in practical expert systems.

B. Medical Relevance and Collaboration

Much of our research has been and will continue to be directly focused on medical AI problems. GALEN, our experimental expert system in pediatric cardiology, is achieving expert levels of performance. Dr. Connelly is initiating a project to develop an expert system based platelet transfusion therapy monitoring program. Dr. Spackman is completing a doctoral thesis on the automated acquisition of rule knowledge in medical microbiology.

Some of our research has focused on problems in diagnostic reasoning and expertise in domains other than medicine. However, our experience indicates that principles of expertise and relevant knowledge engineering tools can cut across task domains. GALEN is demonstrably a useful expert system implementation tool designed in the medical diagnostic task domain. Developments from our work in other domains affecting problems such as automated knowledge acquisition through rule induction and reasoning by analogy will have medical relevance.

Collaboration with Dr. James Moller in the Department of Pediatrics, Dr. Donald Connelly in the Department of Laboratory Medicine, at the University of Minnesota. Dr. Connelly has become a SUMEX user and is teaching a course in medical informatics. He has also initiated a project to create an expert system in platelet transfusion therapy. Collaboration with Dr. Eugene Rich and Dr. Terry Crowson at St. Paul Ramsey Medical Center. Dr. Kent Spackman is a post-doctoral fellow in medical informatics who is completing a Ph.D. thesis in Artificial Intelligence. Dr. Spackman is a resident at the University of Minnesota Hospitals and collaborates with the SOLVER project.

C. Highlights of Research Progress

Accomplishments of This Past Year -- Prior research at Minnesota on expertise in diagnosis of congenital heart disease has resulted in a theory of diagnosis and an embodiment of that theory in the form of a computer simulation model, *Galen*, which diagnoses cases of congenital heart disease [Thompson, Johnson & Moen, 1983]. Continuing development and research with GALEN have led to results in analyzing Garden Path problems in medical diagnosis. Such problems are ones in which an initial solution is later proved to be incorrect. Successful solution of such problems depends upon rejecting an initial incorrect response in favor of a later appropriate one. Errors in Garden Path Problems are generally not due to a lack of knowledge but rather to a confusion over the conditions under which specific rules apply. GALEN was used to identify and test strategies for avoiding Garden Path errors as well as the specific clinical knowledge needed to overcome Garden Path errors in diagnostic reasoning. [Johnson, Moen, and Thompson, 1985].

Galen is descended from two earlier programs written here at Minnesota: *Diagnoser* and *Deducer* [Swanson, 1977]. Deducer is a program that builds hemodynamic models of the circulatory system that describe specific diseases. The models are built by using knowledge about how idealized parts of the circulatory system are causally related. Diagnoser is a recognition-driven program that performs diagnoses by successively hypothesizing one or more of these models and matching them against patient data. The models that match best are used as the final diagnosis. A series of experiments carried out at Minnesota have shown that Diagnoser/Deducer performs as well (and sometimes better) than expert human cardiologists [Johnson et al., 1981].

Despite their early successes, Diagnoser and Deducer did not have a clear, comprehensible structure that is required for the kind of experiments we wish to perform. Galen was built to remedy this problem, taking advantage of the experience gained in the design of Diagnoser and Deducer. Additional discussion of the structure of GALEN can be found in prior annual reports and in the relevant publications.

To determine the generality of our model of expertise in diagnostic reasoning, we are also investigating domains outside medicine. As with our work in congenital heart disease, we have concentrated on the design of mechanisms for structuring problem specific knowledge and for focusing limited computational resources.

One of the Principal Investigators has published results of a study in Expertise in Trial Advocacy, discussing the significance of current research in expertise in legal problem-solving. [Johnson, Johnson, and Little, 1985] Research on legal expertise in corporate acquisition problems has also been investigated. The results of that research suggest that expert corporate acquisition attorneys differ from novices in their greater reliance on internalized norms, prototypes and heuristics. Both expert and novice attorneys in the study went beyond the information provided in task cues in interpreting and predicting actions and situation scripts in the simulated problems. The subjects reasoned heuristically as well as logically. Differences between attorneys in different specialty areas were not large suggesting that the subjects within a domain of problem solving such as legal reasoning acquire meta level reasoning skills that apply to issues within and outside their areas of specialization.

Research is also being completed in a study of cognitive strategies used in making strategic decisions in business. Corporate acquisitions were again used as the context in which to examine expertise. Twenty-four executive subjects were asked to perform an experimental task in which they evaluate companies as candidates for acquisition. The goals of the research are to test for the existence of specialty-related reasoning strategies and to determine the importance of strategic and financial information in problem formulation, problem structuring and choice of strategies in problem solving.

Research in Progress --

Since human experts are notoriously poor at describing their own knowledge, our work requires the creation of problem solving tasks through which experts can reveal criteria for initiating specific hypotheses and methods for investigating those hypotheses.

Current techniques of representing hypotheses and their expectations for diagnosis do not, however, provide much detailed information about the control processes experts use to guide their reasoning. Such control processes typically incorporate highly refined heuristics about which the experts are almost wholly unaware. New research is being proposed to investigate these control structures in legal reasoning, specifically in reasoning by analogy in appellate decision making. Reasoning by analogy appears to be an important inference tool used by experts in many domains as a fundamental problem solving tool. The ability to form plausible analogies lies at the heart of much of the expert ability to be generative when faced with unfamiliar problems. This research will include the implementation of a cognitive simulation of the reasoning by analogy process based upon data obtained by observation of experts solving problems. The results of the simulation will be validated by comparison with human subject data.

We are also investigating several research questions relevant to the architecture of Galen. We have designed an interface to Galen so that users who are unfamiliar with the inner workings of the program can interactively enter case data. Designing the interface raised questions about what forms of data are necessary to adequately and completely represent all possible cases.

One project to test the extensibility of GALEN into other domains is being conducted by a graduate student in the Graduate School of Management. His thesis, Auditing Internal Controls: A computational model of the review process, includes the construction of a working expert system using GALEN. The objective of this study is to formulate and test a model of the processes employed by audit managers and partners in reviewing and evaluating internal accounting controls.

Another project explores the extension of the GALEN architecture into a problem in plant pathology. The main purpose of this research is to find out how the basic postulates about expert reasoning made in Galen hold in a second diagnostic domain. The problem domain chosen for this purpose is Plant Pathology. In collaboration with Professor Paul Teng of the Plant Pathology Department of the University of Minnesota a prototype knowledge base has been implemented. Currently, the knowledge base can diagnose ten potato diseases and has 124 rules. The system is going through evaluation and fine tuning to bring it up to an expert performance level. This system will be useful in the Extension Service at the Plant Pathology department at the University of Minnesota, which provides diagnostic information to farmers over the phone lines.

Dr. Spackman's thesis is entitled "Induction of classification rules under the guidance of comprehensibility-enhancing logical structures and diagnostic performance goals." The purpose of this research is to study and implement methodologies for the automated generation of comprehensible decision rules from empiric data, with emphasis upon logic-based knowledge representation formats and upon problems drawn from the domain of medicine. This work builds upon some of the machine learning methodologies developed at the University of Illinois by R. S. Michalski and others.

This work addresses two shortcomings of previous work on induction of classification rules. These are, first, lack of comprehensibility of the induced rules, and second, lack of flexibility in specifying the diagnostic performance (sensitivity, specificity, or efficiency) desired for the rules that are to be derived.

Comprehensibility of the derived rules or descriptions can be enhanced by imposing restrictions upon the format which the rules may take. For example, the restriction of rules to a unate boolean function format allows the induction of rules that can often be simplified to a "criteria table" type of representation. The type of diagnostic performance a rule must have will depend upon its purpose, and specifying the purpose may allow inductive inference algorithms to trade off small decrements in diagnostic performance for large increments in comprehensibility, or to increase their robustness in the face of noisy or uncertain data.

Successful development of these techniques will lead to enhanced capabilities for deriving rule bases for expert classification systems from empiric data, and will provide new methods for the conceptual analysis of data.

Preliminary results have been obtained for the problem of deriving rules for the identification of bacteria based upon their biochemical profiles in the medical microbiology lab. Other problem domains under investigation are the analysis and interpretation of endocrine laboratory tests, and the induction of rules for the diagnosis of congenital heart disease, for comparison with the rules used in GALEN.

Research is also under way in methods of automating knowledge acquisition in pediatric cardiology. This is being done as thesis research by Paul Krueger. The objective of the research is to design, implement, and test a computerized procedure to derive from examples a nonmonotonic set of rules for an expert classification system. Systems using such rules are generally more efficient than those using monotonic classification processes and more closely approximate psychological models as well.

The research proposes a process for automated learning of preliminary rulebases subject to a set of efficiency constraints which are consistent with a formally defined, psychologically plausible model of classification. The constraints include an upper bound on the amount of information required to explain observations not accounted for by the current set of beliefs, and a lower bound on the degree of inconsistency allowed in the knowledge base at any given time. It will be shown that these constraints can be used to guide the automated determination of both the content and organization of the rules of expert classification systems. The result is behavior that is more focused and efficient, and more closely duplicates the lines of reasoning of domain experts.

A representational formalism for classification knowledge bases based upon a nonmonotonic logic of belief called "autoepistemic logic" (Moore, 1985) is proposed. Having thus defined a representation for the knowledge base the research will propose a methodology for instantiating its concepts within a given application domain. The general approach is to use heuristics to identify from a set of input examples various contextual situations that occur and the types of rules to associate with them. The rule acquisition module (RAM) is then tested in two different application domains. The resulting expert systems will be evaluated for correctness of classification and similarity of their lines of reasoning with those of human experts.

The major conclusion of the research is that constraints similar to those observed in expert human classification processes can be used to guide the empirical induction of efficient expert system rulebases. Supporting this conclusion is the elucidation of a formal nonmonotonic model of classification, and the design and subsequent testing of the Rule Acquisition Module and expert systems derived by it.

D. List of Relevant Publications

1. Connelly, D. and Johnson, P.E.: *Medical problem solving.* Human Pathology, 11(5):412-419, 1980.

2. Elstein, A., Gorry, A., Johnson, P. and Kassirer, J.: *Proposed Research Efforts.* IN D.C. Connelly, E. Benson and D. Burke (Eds.), CLINICAL DECISION MAKING AND LABORATORY USE. University of Minnesota Press, 1982, pp. 327-334.

3. Feltovich, P.J.: *Knowledge based components of expertise in medical diagnosis.* Learning Research and Development Center Technical Report PDS-2, University of Pittsburgh, September, 1981.

4. Feltovich, P.J., Johnson, P.E., Moller, J.H. and Swanson, D.B.: *The Role and Development of Medical Knowledge in Diagnostic Expertise.* IN W. Clancey and E.H. Shortliffe (Eds.), READINGS IN MEDICAL AI, Addison-Wesley, 1984, pp. 275-319.

5. Johnson, P.E.: *Problem Solving.* IN ENCYCLOPEDIA OF SCIENCE AND TECHNOLOGY, McGraw-Hill (in press).

6. Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis.* IN Bloc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

7. 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.

8. Johnson, P.E.: *What kind of expert should a system be?* J. Medicine and Philosophy, 8:77-97, 1983.

9. 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.

10. 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.

11. 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.

12. 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.

13. 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.

14. Johnson, P.E., Hassebrock, F. and Moller, J.H.: *Multimethod study of clinical judgement.* Organizational Behavior and Human Performance 30:201-230, 1982.

15. 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.

16. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Knowledge-based fault localization in debugging.* The Journal of Systems and Software, vol. 3, no. 4 (Dec 83) pp. 301-307, Elsevier.

17. Sedlmeyer, R.L., Thompson, W.B. and Johnson, P.E.: *Diagnostic reasoning in software fault localization.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

18. Smith, K.A., Farm, B., Johnson, P.E.: *Surface: A prototype expert system for selecting surface analysis techniques.* Proceedings of IEEE Conference on Computers and Comm., 1985.

19. Swanson, D.B.: *Computer simulation of expert problem solving in medical diagnosis.* Unpublished Ph.D. dissertation, University of Minnesota, 1978.

20. 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.

21. Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning.* Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

Work in medical diagnosis is carried out with the cooperation of faculty and students in the University of Minnesota Medical School and St. Paul Ramsey Medical Center.

B. Sharing and Interactions with Other SUMEX-AIM Projects

William Clancey, Stanford University, acted as a reviewer of the MEIS Intelligent Systems Project in September, 1984 at the University of Minnesota. The Principal Investigators in the SOLVER project are also principal investigators in that project.

Paul Johnson was a panel member at the SUMEX-AIM conference in Columbus, Ohio in 1984. Dr. Connelly and two graduate students associated with the SOLVER PROJECT also attended the conference.

III. RESEARCH PLANS

A. Project Goals and Plans

Near term -- Our research objectives in the near term can be divided in three parts. First, we are committed to the design, implementation, and evaluation of Galen, as described above. We have completed an interactive front end so that physicians can directly enter patient data, and Galen's knowledge base is currently being "tuned" with the help of Dr. James Moller, an expert physician collaborator from the University of Minnesota Pediatric Cardiology Clinic, the Diagnoser program, and with expert physicians. We believe that GALEN has passed through phases of expertise assessment and cognitive simulation and that it is now approaching a level of performance that will qualify it as a true expert system. An objective now is to extend the explanation capability of GALEN. We are initiating a new investigation into two aspects of expert problem solving that relate to the interaction between a problem solving system and its environment: *"query generation"* and *explanation*. Some simple expert systems proceed from a fixed set of input data to an evaluation of that data. For most problem domains, however, the space of possibly relevant information is large, and some or all of this information may have costs associated with its acquisition. Thus, computational and other costs can be reduced by some mechanism which intelligently selects appropriate queries designed to solicit information that is relevant and cost effective in terms of the problem being solved. Expert systems for complex problem domains must also be able to generate explanations for their actions. Unless the system operates in an entirely autonomous manner, users must be apprised of the rationale for system actions. There is a particular need for explanations tailored for system users rather than system designers.

Experienced experts are typically quite proficient at asking relevant questions, even when the criteria for relevance is difficult to specify. These experts use heuristics capable of keying on selected aspects of data already examined and on the current problem state in order to select the next needed query. We propose to incorporate these heuristics into a *"query generation knowledge base"* . This knowledge base can be thought of as a form of domain specific meta-knowledge. It contains rules by which the problem state can be efficiently evaluated in order to determine the next course of action. By basing these rules on actual expert knowledge and experience, it will often be possible to bypass the combinatorial complexity associated with either blind search or optimization techniques.

Our approach to explanation starts from the premise that substantially different forms of explanation are required within a single expert system. The type of explanation is distinguished both by the level of sophistication of the person receiving the explanation and by whether that person is principally interested in the specific problem being solved or in the internal working of the expert system. Less sophisticated users of the system are likely to have only a superficial understanding of the nature of the system being diagnosed and will require explanations in terms of simplified system properties with which they are familiar. Expert users will require information about significant details of the state of the system being diagnosed and the causal relationships that connect system state with observable symptoms. Designers and maintainers of the expert system require explanations in terms of the actual lines of reasoning used to arrive at a decision.

We will be focusing principally on providing explanations for system *users* rather than system *designers*. Explanations for users must be phrased in terms of the system being diagnosed. Descriptions of the system itself are more important that descriptions of the reasoning strategies used to understand the system. For example, many diagnostic tasks are efficiently approached utilizing recognition-based reasoning strategies using knowledge arising from empirical association. Experts (or possibly automatic learning systems) learn to associate particular interpretations with particular patterns in the data. For many problem domains, knowledge of this sort is quite powerful, providing accuracy without the complexity associated with causal reasoning. The user of such a system, however, requires explanations in terms of causality. This suggest a two-step process. Problem solving is done using a recognition-based strategy. Explanations are generated by combining the results of this process with additional, causally-based explanation knowledge.

Our second objective consists of making extensions to the knowledge capturing strategies developed in our original work in medical diagnosis. In the near term this work will examine descriptive strategies in which experts attempt to use a formalized language to express what they know (e.g. production rules), observational strategies in which experts perform tasks designed to reveal information from which a theory of task specific expertise can be built, and intuitive strategies in which either experts behave as knowledge engineers or knowledge engineers attempt to perform as pseudo experts. The research projects of Dr. Spackman and Paul Krueger which have been discussed previously are both directed toward this objective.

Our third near term objective will be to investigate one of the central problems of recognition based problem solving, how to classify problems when solving them. Questions related to problem classification which we will be examining include: What patterns do experts and novices detect in a problem that allows them to classify it as an instance of a problem type that is already known? How does an expert make an initial choice of the level of abstraction to be used in solving a problem? How can an expert recover from an initial incorrect choice of levels? How can the difference between causal and prototypic modes of reasoning be modeled as differences in levels of abstraction, and how can a common model for these two types of reasoning be constructed? We will be pursuing these questions in the areas of problem solving like law, auditing, and management, as well as in medicine.

Long range -- Our long range objective is to improve the methodology of the "knowledge capturing" process that occurs in the early stages of the development of expert systems when problem decomposition and solution strategies are being specified. Several related questions of interest include: What are the performance consequences of different approaches, how can these consequences be evaluated, and what tools can assist in making the best choice? How can organizations be determined which not only perform well, but are structured so as to facilitate knowledge acquisition from human experts? In the coming year we will be exploring these questions in areas of design and management as well as in law, management and medicine.

B. Justification and Requirements for Continued SUMEX Use

Our current model development takes advantage of the sophisticated Lisp programming environment on SUMEX. Although much current work with Galen is done using a version running on a local VAX 11/780, we continue to benefit from the interaction with other researchers facilitated by the SUMEX system. We expect to use SUMEX to allow other groups access to the Galen program. We also plan to continue use of the knowledge engineering tools available on SUMEX.

We are working toward a Commonlisp implementation of the GALEN system and expect to rely heavily on Commonlisp for future projects.

One of our students implemented a demonstration legal expert system in EMYCIN using the SUMEX resource, and we still find that the resource is valuable for making available major systems which we do not have locally, such as EMYCIN.

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

Our current grant from MEIS has permitted us to purchase four Perq 2 AI 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.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. CAMDA Project

CAMDA Project

CAMDA Research Staff:

Prof. Samuel Holtzman, Co-PI	Engineering-Economic Systems
Prof. Ronald A. Howard, Co-PI	Engineering-Economic Systems
Prof. Ross Shachter	Engineering-Economic Systems
Leonard Bertrand	Engineering-Economic Systems
Jack Breese	Engineering-Economic Systems
Kazuo Ezawa	Engineering-Economic Systems
Keh-Shiou Leu	Engineering-Economic Systems
Seok Hui Ng	Engineering-Economic Systems
Emilio Navarro	Engineering-Economic Systems
Dr. Adam Seiver	Engineering-Economic Systems
Joseph Tatman	Engineering-Economic Systems
Dr. Emmet Lamb	School of Medicine
Dr. Robert Kessler	School of Medicine
Dr. Frank Polansky	School of Medicine

Associated faculty:

Prof. Edison Tse	Engineering-Economic Systems

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The Computer-Aided Medical Decision Analysis (CAMDA) project is an attempt to develop intelligent medical decision systems by combining the descriptive generality of expert-system technology with the normative power of decision analysis.

B. Medical Relevance and Collaboration

The primary effort of the CAMDA project during 1984 and early 1985 has been focused on the design and implementation of RACHEL, an intelligent decision system for infertile couples. This system is designed to help patients and physicians deal with difficult medical treatment choices. RACHEL is being developed in close cooperation with the Engineering-Economic Systems Department, the Obstetrics and Gynecology Department, and the Surgery Department (Urology Division), all at Stanford.

In addition to the development of RACHEL, there are several active research programs within the CAMDA project. One such program is aimed at developing a representation for dynamic decision processes (such as those faced by cancer patients) that do not necessarily satisfy the Markov assumption. Another is concentrating on the development of fast algorithms for the solution of general decision problems.

A recent addition to our research project is a program to design cost-effective strategies for monitoring the recurrence of bladder cancer.

C. Highlights of Research Progress

C.1 Accomplishments this past year

We have successfully implemented a pilot-level version of RACHEL. As we define it, a pilot system is one where the essential algorithms work individually as well as interactively with one another, operating with knowledge that is representative of the system's domain. Such a system lacks two important elements that must exist within a prototype-level implementation: an extensive knowledge base, and a front end usable by trained users who may not be familiar with the details of the system.

As part of the development of RACHEL, we have developed a facility to construct individualized models of the patient's preferences over the set of possible outcomes of an infertility therapy. This facility operates in two consecutive stages. The first stage constructs a parametric model from a library of plausible model elements. A typical consideration at this stage is whether to explicitly account for the patient's lifetime. For instance, a treatment strategy which involves surgery would warrant such explicit consideration, whereas a therapy consisting strictly of drugs would not. The second stage in the preference model development process involves the assessment of specific parametric values. These values are obtained directly from the patient to ensure that the overall preference model genuinely reflects his or her desires.

It is important to note that since the preference model is built to fit the specific needs of each case, the interaction between the patient and the system is short and well-focused. In particular, the patient is only asked to respond to a few (about five to ten) questions. These questions are selected so that their relevance to the case is intuitively obvious from the patient's point of view.

Also as a part of RACHEL, we have developed a knowledge base dealing with the decisions faced by the subset of infertile couples whose inability to conceive has been traced to a blockage of the Fallopian tubes of the female partner. In particular, the knowledge in RACHEL deals with the choice between two important procedures pertinent to this condition: laparotomy and in-vitro fertilization.

Another accomplishment during this past research year has been the improvement of our influence-diagram solution procedure. In its original form, this procedure essentially took a brute-force approach to the solution of well-formed influence diagrams. Although its solutions were mathematically correct, the program was inefficient in terms of both computational time and storage requirements. In its current implementation, the program is considerably more efficient and has an adequate front end which makes it accessible to a fairly wide class of users. Empirical results indicate that the size and complexity of problems that can be represented and solved with the system not only exceed the bounds of its original design, but are comparable and possibly superior to those of the best commercially available decision-analytic software.

Similarly, RACHEL's inference engine has been improved in several important ways. Prominent among these are a means for attaching general procedures at any point in the inference process, a variety of built-in procedures for the acquisition and display of information coupled with a facility for controlling these procedures (i.e., for the control of ASKability and TELLability), and a simple explanation mechanism.

C.2 Research in progress

The RACHEL system continues to be developed along four distinct directions: the efficiency and flexibility of RACHEL's inference engine are being improved, its explanation mechanism is being enhanced, RACHEL's facility for the development of patient preference models is being upgraded, and its knowledge base is being enlarged.

As it is currently implemented, the inference engine used by RACHEL is quite inefficient. This inefficiency is, to some extent, a deliberate design choice since the engine was designed to be very general and highly modular. Thus, there are many procedural redundancies and much unnecessary baggage in the programs that implement it. Now that we have a clearer idea of how the engine is to be used we have redesigned it by doing away with some of the original generality and modularity in favor of a more efficient process. Furthermore, the new design emphasizes and enhances particularly useful engine features such as its ASKability and its TELLability.

A further enhancement to RACHEL's inference engine concentrates on the system's ability to explain its line of reasoning. The original design only responds to online "why" queries by displaying its dynamic goal stack. In its new form, the engine allows offline as well as online queries in both "why" and "how" formats.

Beyond traditional explanation capabilities, we are exploring possible means to explain decision-theoretic inferences. In particular, we are trying to understand how to explain decision recommendations that are based on the maximization of expected utility to users unfamiliar with decision theory. Our current research indicates that a promising way to do this is to break down large decision problems into smaller, more manageable pieces whose formal solution can be checked against intuition. Although still at an early stage, this line of research seems to be on the path of eliminating an important barrier to the widespread use of normative decision techniques.

An exciting area of current interest is the improvement of RACHEL's facility for the creation and assessment of parametric models of patient preferences. In particular, we are trying to increase the generality of RACHEL's model library to account for acute as well as chronic conditions and to simplify the corresponding assessment process. This simplification is based on the notion that a better understanding of the major concerns of patients can help us redesign the questions asked by RACHEL so that they are closer to the specific experiences of individual patients. As part of this effort, we expect to have significant contact with actual patients to ensure the clinical relevance of our research.

A fourth area where RACHEL is being enhanced is the expansion of its medical and decision-analytic knowledge bases. Planned additions include further knowledge about the treatment of tubal blockage (including more data on in-vitro fertilization procedures and an ability to consider a wider class of patients) and a new packet of knowledge dealing with deterministic sensitivity analysis.

In addition to the development of RACHEL, there are several active research programs within the CAMDA project. One such program is aimed at developing a representation for dynamic decision processes (such as those faced by cancer patients) that do not necessarily satisfy the Markov assumption. This research has led to a generalization of influence diagrams which allows multiple value nodes. This generalization makes it possible for complex sequential decision processes (whose solution would otherwise be infeasible) to be efficiently solved.

Another research program within the CAMDA project is the development of fast algorithms for the solution of decision problems formulated as influence diagrams. In general, the solution of an influence diagram (i.e., the calculation of a recommended decision strategy) is obtained by the repeated application of an operation, known as "removal", to all nodes in the diagram other than the value node. The removal of a node in the diagram is a generalization of the foldback operation needed to solve a decision tree. With rare exceptions, the order in which nodes are removed from a diagram is not unique. Current results indicate that significant reductions in the computational burden of solution can be achieved by controlling the order in which diagram nodes are selected for removal.

At a more fundamental level, we are exploring the consolidation of the predicate calculus with probabilistic logic. Of particular interest is the design of an integrated inference engine that performs logical inferences within a probabilistic framework. A central problem in this research is the definition of universal and existential quantification in probabilistic terms.

A recent addition to our research project is a program to design cost-effective strategies for monitoring the recurrence of bladder cancer. We expect this research to interact with our ongoing search for more effective models of patient preferences.

D. Publications

1. Holtzman, S.:*A Model of the Decision Analysis Process*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1981.

2. Holtzman, S.:*A Decision Aid for Patients with End-Stage Renal Disease*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

3. Holtzman, S.:*On the Use of Formal Models in Decision Making*, Proc. TIMS/ORSA Joint Nat. Mtg., San Francisco, May, 1984.

4. (*) Holtzman, S.: *Intelligent Decision Systems*, Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1985.

5. Shachter, R.: *Evaluating Influence Diagrams*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1984.

6. Shachter, R.: *Automating Probabilistic Inference*, Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1984.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

II.A Medical Collaborations and Program Dissemination Via SUMEX

Since its inception, the CAMDA project has benefited from an active relationship among decision analysts, computer scientists, and members of the Stanford medical community. In particular, RACHEL is being developed in close cooperation with physicians in the Infertility Clinic at Stanford. Other programs within the CAMDA project such as our research on the form and use of medical preference models are being done in cooperation with physicians at the Palo Alto Veterans Administration Hospital and at El Camino Hospital.

II.B. Sharing and Interactions with other SUMEX-AIM Projects

II.B.1 SUMEX-AIM 1984 Workshop:

Samuel Holtzman participated in the 1984 AIM workshop in Columbus, Ohio. In addition to the presentation of a summary of CAMDA research, he had many opportunities to interact with workshop participants on an informal basis. Of particular interest were several discussions with members of the MIT/TUFTS group interested in medical decision analysis which have led to an interchange of ideas that continues to this date.

II.B.2 Decision Systems Laboratory Research Meetings

As part of the CAMDA project, we have instituted a weekly research meeting for those interested in the design and implementation of computer-based decision systems. This weekly meeting has become a very active forum for the presentation of research results. The following topics of direct relevance to medical decision making were presented during the last two academic quarters.

Date	Speaker	Topic
03-OCT-84	Ross Shachter	Probabilistic Inference
17-OCT-84	Jack Breese	Dempster-Shafer Theory
24-OCT-84	Kazuo Ezawa	Efficiency in Solving Influence Diagrams
07-NOV-84	Majid Khorram	Fuzzy Sets and Decision Making
14-NOV-84	Dan Kent	Utility Theory Underlying Physicians' Treatment Thresholds: HELP!
21-NOV-84	Yann Bonduelle	Explanation in Decision Systems
09-JAN-85	Ross Shachter	What Do You Call the Offspring of SUPERID and INFLUENCE?
23-JAN-85	Doug Logan	The Value of Probability Assessment
06-FEB-85	Seok Hui Ng	Minimal Tumor Follow-up Examination Schedule for Recurrent Bladder Cancer Patients.
13-FEB-85	Keh-Shiou Leu	TEREISIAS' Explanation Facility
06-MAR-85	Joe Tatman	Algorithm for Decision Processes Optimization
13-MAR-85	Gerald Liu (UC)	Knowledge Structure in Evidential Reasoning

II.B.3 Course in Medical Decision Analysis

A new course in medical decision analysis, taught by Prof. Samuel Holtzman, is being offered for the first time during the Spring quarter of 1985. The course is offered jointly by the Engineering-Economic Systems Department, the Medical Information

Sciences Program, and the Computer Science Department. The objective of the course is to expose students to the practice of decision analysis for clinical purposes and to introduce them to the design and use of computer-based medical decision tools.

II.C. Critique of Resource Management

The CAMDA project is heavily dependent upon the availability of the SUMEX computing resource. The physical facility as well as the staff of SUMEX-AIM are excellent. In particular, it has been a pleasure to deal with Ed Pattermann, who is invariably courteous, responsive to our needs, and effective in his actions. We will certainly miss him now that he has moved to industry. Pam Ryalls has also provided much needed help in managing the CAMDA project in a manner that is friendly and efficient.

As an update to last year's report, the previously reported Ethernet deficiencies have been corrected. This improvement was part of a campus-wide effort to improve Stanford's computer network which directly affected our campus connection to SUMEX. The system load on SUMEX continues to be heavy, although it appears to be somewhat lower than it was last year. The ability of the CAMDA project to use the DECSYSTEM-2020 machine operated by SUMEX (referred to as TINY) has had a significant effect on our ability to demonstrate our systems during normal business hours, further reducing our frustration with the main system's load.

III. RESEARCH PLANS

III.A Project Goals and Plans

During the upcoming year, we intend to enhance four specific elements of the RACHEL system: its inference mechanism, its explanation facility, its ability to model patient preferences, and its medical and decision-analytic knowledge bases. Furthermore, we intend to continue to improve our understanding of normative decision methodologies, with particular emphasis on the use of these methodologies for computer-based decision support. Section I.C.2 describes the near-term goals of the CAMDA project in more detail. Our long-term goal remains that of designing and implementing usable, fully-validated and documented systems for medical decision support.

III.B Justification and Requirements for Continued SUMEX Use

The CAMDA project is truly interdisciplinary. It draws on elements of decision analysis, artificial intelligence, and medical science. The project has the potential to contribute to each of these disciplines in important ways.

In particular, the CAMDA project is likely to lead to the development of tools and techniques that greatly improve the quality of decision making in medicine. For instance, RACHEL explicitly considers uncertainty, decision alternatives, and patient preferences in developing recommendations. In spite of its generality, RACHEL's interaction with the user is sufficiently terse and simple to support the claim that systems based on its methodology can be effective clinical decision tools. Much of the simplicity and terseness of RACHEL's operation is a direct consequence of the AI foundations of the system's design.

The heavy reliance of the CAMDA effort on artificial intelligence technology make SUMEX-AIM an ideal environment in which to pursue this research.

III.C Needs and Plans for other Computing Resources beyond SUMEX-AIM

The CAMDA project has access to four Olivetti M24 and one MAD-1 personal computers (IBM-PC type) as well as to one Apple Macintosh (128K) computer. In addition, we continue to search for funds to acquire one or more state-of-the-art LISP machines.

III.D Recommendations for Future Community and Resource Development

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

A major benefit provided by the existing SUMEX-AIM facility is the availability of very low-cost computing resources. Access to these resources is granted primarily on the basis of an assessment of the value of the proposed research to the overall goal of making artificial intelligence a useful medical tool. Imposing fees for using SUMEX would prevent users with modest means from obtaining access to the facility on the basis of merit alone.

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

The CAMDA project has access to several personal computers for its research. These machines include Olivetti M24's (marketed as the A.T.&T. personal computer in the U.S.) and a MAD-1 personal computer -- all of which are compatible with the IBM-PC. In addition, the project has purchased an Apple Macintosh. These machines are used as a supplement to the SUMEX mainframe, and are not intended to replace it.

IV.C.2. Protein Secondary Structure Project

Protein Secondary Structure Project

Robert M. Abarbanel, M.D., Ph.D.
Section on Medical Information Science
University of California Medical Center
University of California at San Francisco

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Development of a protein structure knowledge base and tools for manipulation of that knowledge to aid in the investigation of new structures. System to include cooperating knowledge sources that work under the guidance of other system drivers to find solutions to protein structure problems. Evaluations of structure predictions using known proteins and other user feedbacks available to aid user in developing new methods of prediction.

B. Medical Relevance and Collaboration

Many important proteins have been sequenced but have not, as yet, had their secondary or tertiary structures revealed. The systems developed here would aid medical scientists in the search for particular configurations, for example, around the active sites in enzymes. Predictions of secondary structure will aid in the determination of the full "natural" configuration of important biological materials. Development of systems such as these will contribute to our knowledge of medical scientific data representation and retrieval.

C. Highlights of Research Progress

The prediction of beta-alpha protein structures was completed in 1982. The system was developed on a VAX 11/750 at the University of California, San Francisco, to allow researchers to describe patterns of amino acid residues that will be sought in the sequences under study. The presence or absence of these "primary" patterns was then combined with other measures of structure, like hydrophobicity, to suggest possible alpha helix or beta sheet or turn configurations.

The segments of a sequence between turns were then analyzed to determine the allowable extent of the possible secondary structure assignments. Any segments remaining were then used to generate all possible complete structures. Only two beta strands with the character of sheet edges are allowed in any prediction. This hierarchical generation and pruning resulted in nearly 95% turn prediction accuracy, and excellent delimiting of helices and sheets. In some cases, one and only one secondary structure was predicted.

Work in progress -- The original pattern matching and manipulation system written in the C language, was re-written in Franz Lisp to run under UNIX(TM). This system was then re-written to run under KEE, The Knowledge Engineering Environment (TM, Intellicorp) on a Symbolics 3600 at the Computer Graphics Laboratory, University of California, San Francisco. This environment provides for ready display of pattern matches and viewing and manipulating of the applications of sets of rules. The original α/β rules are being tested now and new sets of rules are under development for classifications of unknown structures and the assignment of turns and secondary labels to regions of those structures.

D. List of Relevant Publications

Cohen, F.E., Abarbanel, R.M., Kuntz, I.D. and Fletterick, R.J.: *Secondary structure assignment for α/β proteins by a combinatorial approach*, Biochemistry, 22, pp 4894-4909, (October 1983).

At this time, another paper on prediction of "turns" in several classes of proteins has been accepted by Biochemistry for publication.

Abarbanel, R.M., Wieneke, P.R., Mansfield, E., Jaffe, D.A., Brutlag, D.L.: *Rapid searches for complex patterns in biological molecules*, Nucleic Acids Research, 12, pp 263-280, (January 1984).

Abarbanel, R.M.: *Protein Structural Knowledge Engineering, Ph.D. thesis*, University of California San Francisco, (December, 1984).

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

None.

B. Sharing and Interactions with SUMEX Projects

This project is closely allied with the MOLGEN group, both in computer and scientific interests. Some pattern matching methodology created for the protein data base has been adopted and used in the various DNA knowledge bases. The principal persons in the MOLGEN group have contributed to this project's use and understanding of knowledge base software and resources.

C. Critique of Resource Management

Work continues on the UNIX systems at the University of California, San Francisco and on the Symbolics Lisp Machine there. SUMEX has been used primarily for communications with other researchers.

Resource management remains excellent. The staff are friendly and responsive. Network access, bulletin boards and the mail system have provided a means to collaborate with others doing related work locally as well as in Europe. SUMEX-AIM staff have been most helpful in getting this project started on the Dolphin workstations and in providing an environment where new tools have been made available for use.

III. RESEARCH PLANS

A. Project Goals and Plans

Since the funding for this project has been terminated, remaining work will be supported by Prof. I. Kuntz at UCSF. Development of the KEE based pattern matching and structure inference system continues.

In particular, at this time, an improved general sequence pattern matching facility has been implemented. A hierarchy of pattern types has been developed so that each pattern may inherit methods for evaluation and display, from common ancestor units. Evaluation of patterns and collections of patterns on the 3600 is from 4 to 10 times faster than under Franz lisp on the Vax/750 running UNIX. Display of matches has been made interactive so that the sequence is shown with mouse sensitive regions and pattern symbols allowing a user to determine the reasons for a match. This feedback allows for improved pattern design.

A KEE TellAndAsk operator is being developed that will allow the rule system to interact with the pattern matchers thus allowing inference about patterns and the suggested underlying structure.

Work will continue on this project though at a slow pace due to the other commitments of the principal investigator. As other resources become available, it is hoped that new rule sets may be developed and tested during the next project year.

B. Need for Resources

-- no comment

C. Recommendations

-- no comment

IV.C.3. REFEREE Project

REFEREE Project

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

Byron W. Brown, Ph.D.
Dept. of Biostatistics
Stanford University

Daniel E. Feldman, Ph.D., M.D.
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The goal of this project is two-fold: (a) use existing AI methods to implement an expert system that can critique medical journal articles on clinical trials, and (b) in the long term, develop new AI methods that extract new medical knowledge from the clinical trials literature. In order to accomplish (a) we are building the system in three stages.

1. System I will assist in the evaluation of the quality of a single clinical trial. The user will be imagined to be the editor of a journal reviewing a manuscript for publication, but the program will be tested on a variety of readers, including clinicians, medical scientists, medical and graduate students, and clerical help.

2. System II will assist in the evaluation of the effectiveness of the treatment or intervention examined in a single published clinical trial. The user will be imagined to be a clinician interested in judging the efficacy of the treatment being tested in the trial.

3. System III will assist in the evaluation of the effectiveness of a single treatment examined in a number of published clinical trials.

B. Medical Relevance

The burden of "keeping up with the literature" is particularly onerous in the practice of medicine and in medical research [30, 31]. Reading the abstracts in a few journals and selecting several key articles for a rapid survey are the best that most clinicians can hope to accomplish each week. The time and effort necessary for a thorough and critical reading of even a few research reports are not available.[1] Sackett reports that to keep up with the 10 leading journals in internal medicine a clinician must read 200 articles and 70 editorials per month [31]. It was also estimated that the biomedical

[1] In an informal check on this intuition two of us, with considerable training in analyzing clinical trials (BWB and DEF) timed critical readings of a five page article on a clinical trial in the New England Journal of Medicine [3]. Our times were 30 and 120 minutes.

literature is expanding at a compound rate of 6% to 7% per year, or doubling every 10 - 15 years [31, 28]. Furthermore, even if more time were available the statistical and epidemiological skills necessary for critical reading are not part of most clinicians' repertoires[1]; and yet decisions about which therapy to use, what intervention to adopt, or what advice to give patients must be based on a combination of clinical experience and published literature. But the existing literature is often confusing and contradictory [20]and publication in the most prestigious medical journals does not guarantee freedom from serious methodologic flaws and erroneous conclusions [22, 8]. Any assistance to the clinician must deal with both the problem of the vastness of the literature and the quality of the research report. Similar problems are faced by the editors of medical journals, swamped with manuscripts to review and evaluate, and by research scientists and academicians trying to stay abreast of the developments in their fields. How can they cover more and yet evaluate better and more consistently? Clearly any machine assistance would be welcome.

C. Highlights of Progress

This project is just getting started.

Preliminary work has been done on REFEREE [10], a prototype expert system for determining the quality of a clinical trial report, and the efficacy of the intervention evaluated in the trial. REFEREE is written in EMYCIN, a rule-based programming language which allows rapid prototyping of a consultation system that gives advice to a user. It presupposes that a knowledge base about the problem area has been constructed, which usually involves codifying an expert's knowledge.

The basic format of a REFEREE session is fairly simple. The reader is asked a series of questions pertaining to the paper and the study described. The answers given are used to rate the overall quality of the paper and the probable efficacy of the treatment described. (See sample dialogs below).

In the first version of REFEREE, after the program has finished with its chain of questions and deductions, the quality of the paper and the efficacy of the drug are given to the user as a "merit score", an integer between 0 and 10, with 10 indicating the highest quality. Additionally, the user is provided with a series of English language messages indicating the main flaws detected in the paper. The merit score was used because the expert system makes its judgements by using a weighted average of values assigned to each aspect of the paper being critiqued. As the user answers the consultant's questions, the answers are given individual merit scores. For example, if the user's answer indicate that experimental blinding was done correctly, the paper is given a high score in the blinding category. When all merit score assignments have been made, the total merit score is calculated as a weighted average of the categorical merit scores, with those categories that are more crucial to a good paper or clinical trial being given a higher weight.

The final result of this calculation is a number between 1 and 10 which serves as a quality measure for the paper or the treatment. A 1 indicates low quality; a 10 indicates the highest quality. An integer as a final result, however, can be very cryptic. It is usually quite difficult, given just an integer, to understand or believe the findings of the consultant. It was discovered quite early that users, when presented with just the bare merit score of the paper, would want to know why the paper was rated in the way it was. For this reason, English language statements are given to the user, indicating the nature of the main flaws of the paper. In each category, if the calculated merit score is

[1]A recent survey of the statistical methods used by authors in the New England Journal of Medicine indicated that 42 per cent of the articles surveyed relied on statistical analysis beyond descriptive statistics [6].

found to be less than an arbitrary minimum, this is noted in a sentence or two, and given to the user at the end of the consultation. In this way, the user not only gets an overall picture of the quality of the paper, but also an indication of the general areas in which the paper was found to be lacking.

Several problems were found in the original version of REFEREE. It was discovered that the use of a weighted average precluded the use of EMYCIN's certainty factors. Because of this, the user would often be forced to choose from a fairly limited set of possible answers to the consultant's questions. The lack of versatility implied by this constraint dictated that a new approach which could make full use of EMYCIN's certainty factors should be used.

In order to do this, the old rule base was scrapped, and a new one was written. Instead of deciding on a rating between one and ten to indicate quality, the new version simply decides whether or not the paper in question is of "high academic and scholarly quality", with an EMYCIN certainty factor modifying the conclusion. For example, in the case of a mediocre paper, the program would conclude that the paper was of "high quality", but only with a certainty of say, .5, on a scale between -1 and 1. Though the words "certainty factor" are used for historical reasons, our final number is the equivalent of a merit score.

While at first glance the two approaches seem similar, the second approach was found to be much more flexible and satisfying from the user's standpoint. Since the conclusion is in terms of the programs certainty that the paper's quality is good, the user may incorporate his or her own uncertainty into the dialogue with the program. This was accomplished by asking mainly yes/no questions, and at all times allowing the user to indicate his or her certainty in the answers given. Thus, if the program asks the user if the quality of the paper's literature review was high, he or she can answer simply "yes" or "no", indicating complete confidence in the answers, or modify a yes/no answer with a certainty factor, indicating that he or she is not completely certain. The user's answers, along with the uncertainty indicated by him or her, will be combined by EMYCIN to give a final conclusion on the paper's quality.

As an example, one of the old-style rules might have been something like this: If the user indicates that the literature review is of "poor quality", conclude that the merit of the paper is 3 with a (built-in) weight of 2. After all the merit values had been calculated, a weighted average, (using built-in weights) would be taken to come to the final merit score. In contrast, one of the new rules would be of the form: If the user gives a "yes" answer to the question "Is the literature review thorough and balanced?", conclude that the paper is of good quality with a certainty of .3. While in the first case the user was limited to a set of possible answers (e.g. excellent, good, poor), the second rule gives the user the opportunity to answer either yes or no, and qualify that answer with any degree of certainty desired. If, in the second rule, the user gives a certainty of less than 1 that the literature review was of good quality, the inferred conclusion about the quality of the paper will be automatically downgraded as well. In other words, if the user expresses uncertainty, the conclusion about the quality of the paper will be less certain.

The new approach, in addition to supplying the user with the ability to express varying degrees of uncertainty, also allows for a hierarchical question structure. At any point, if the user is unclear of the appropriate response, the program can prompt with further, more detailed questions, until a conclusion about the original question can be provided. Conversely, whenever a user is willing to give an answer, the program will refrain from dwelling on the issue and omit its long series of sub-questions. In this manner the amount of detail provided can be individualized.

This current version of REFEREE has two hundred rules and has been tested by the present research team on several papers. It is this program that will be expanded as described in Section III-A. Part of a sample consultation is shown below.

```
   --------MEDICINE-1--------
   The first paper of MEDICINE-1 will be referred to as:

      --------PAPER-1--------

         --------STATISTICS-1--------
         1) What is the size of the control sample?
         ** 25
         2) How many of the subjects in the control sample responded to
            treatment?
         ** 14
         3) What is the size of the test sample?
         ** 23
         4) How many of the subjects in the test sample responded to
            treatment?
         ** 23
...

         --------PLANNING-1--------
         9) Was there an explicit stopping rule defined before the experiment
            was run?
         ** N

         --------RANDOMIZATION-1--------
      10) Was there any mention of the use of randomization in patient
          assignment?
      ** Y
      11) Was the assignment of subjects in the experiment performed blindly?
      ** UNK
...
         --------BLINDING-1--------
      16) Was the experiment double blinded, or was any mention made of
          blinding in the experiment?
      ** Y
      17) Was there any mention of an effort to make the placebo and
          medication as similar as possible?
      ** N
...
The strength of the evidence indicating the efficacy of PAPER-1 is as
     follows:
        There is some evidence for efficacy, but further study is needed.

The general quality of the paper is as follows:
        The current paper is of poor quality.

The flaws of the current paper are as follows:
        A stopping rule was not defined or was not adhered to in the
            experiment.
        The measures taken to evaluate subject compliance were inadequate or
            non-existent.
        Subjects were not randomly assigned treatment groups, seriously
            weakening the validity of the conclusions.
        Though an effort was made to blind the experiment, the techniques
            used were not effective.

The final calculated efficacy of the drug as indicated by the given clinical
     trial (between 0 and 10, with a score of 10 being the highest) is as
     follows:
        5.

The final merit of the current paper is as follows:
        3.

   23) Are there any other papers on MEDICINE-1?
   ** N
   24) Do you want the results of this consultation output to a file?
   ** N
```

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations

Dr. D. Feldman is a physician and epidemiologist at the Stanford Center for Disease Prevention. Prof. B. Brown is currently teaching a Medical School class on reading medical journal articles.

B. Interactions with other SUMEX-AIM projects

Our interactions have all been through the Knowledge Systems Laboratory where we have discussed design and implementation issues.

C. Critique of Resource Management

The SUMEX staff has been most cooperative in helping get this project started. We have tried to place few demands on the SUMEX staff, but have received prompt answers to all questions.

III. RESEARCH PLANS

A. Goals & Plans

It is proposed to construct three computer-based expert systems to assist a variety of different readers in the evaluation of an extensive but well defined area of the medical literature, clinical trials. It is further proposed to test the hypothesis that such programs will enable a variety of users to read the literature on clinical trials more more critically and more rapidly.

The expert systems will be developed using the EMYCIN programming environment and the production rule approach followed successfully in previous expert systems [11, 17, 21, 24, 4].

The three programs to be developed are separate, but closely related:

1. System I will assist in the evaluation of the quality of a single clinical trial. The user will be imagined to be the editor of a journal reviewing a manuscript for publication, but the program will be tested on a variety of readers, including clinicians, medical scientists, medical and graduate students, and clerical help.

2. System II will assist in the evaluation of the effectiveness of the treatment or intervention examined in a single published clinical trial. The user will

be imagined to be a clinician interested in judging the efficacy of the treatment being tested in the trial.

3. System III will assist in the evaluation of the effectiveness of a single treatment examined in a number of published clinical trials.

Within the duration of this research it is also proposed to test the first two systems against unassisted evaluations by the various categories of readers. The testing will include a formal testing of the programs by comparing the speed and number of flaws found in using the program with similar measurements on unassisted reading. In addition there will be a more informal evaluation by questionnaire of the subjective impressions of users of the program, ascertaining the likelihood of routine use and the value of such a program to the user.

This proposal with its concentration on clinical trials is regarded as the initial step in a more general research goal - building computer systems to help the clinician and medical scientist read the medical literature more critically.

B. Justification for continued SUMEX use

We will continue to use SUMEX for developing the AI methods. We need EMYCIN at the moment because it provides a good environment for building a rule-based system that may grow to many hundreds of rules. EMYCIN is not available on other machines without substantial cost.

C. Need for other computing resources

In the short term we will not need additional resources. Should we decide to implement a new system in a framework other than EMYCIN, we might seek funding to buy a LISP workstation.

D. Recommendations

Although our use has been small, we find the load average on SUMEX often precludes running test cases during the day. We have no specific recommendation, but would like to have access to small amounts of high quality computer time.

IV.C.4. Ultrasonic Imaging Project

Ultrasonic Imaging Project

James F. Brinkley, M.D.
W.D. McCallum, M.D.
Depts. Computer Science, Obstetrics and Gynecology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

This report is a summary of the overall accomplishments of the ultrasonic imaging project since it is currently being discontinued. The long range goal of this project was the development of an ultrasonic imaging and display system for three-dimensional modelling of body organs. The models would be used for non-invasive study of anatomic structure and shape as well as for calculation of accurate organ volumes for use in clinical diagnosis. Initially, the system was used to determine fetal volume as an indicator of fetal weight; later it could be adapted to measure left ventricular volume, or liver and kidney volume.

The general method we used was the reconstruction of an organ from a series of ultrasonic cross-sections taken in an arbitrary fashion. A real-time ultrasonic scanner is coupled to a three-dimensional acoustic position locating system so that the three- dimensional orientation of the scan plane is known at all times. During the patient exam a dedicated microcomputer based data acquisition system is used to record a series of scans over the organ being modelled. The scans are recorded on a video tape recorder before being transferred to a video disk. 3D position information is stored on a floppy disk file. In a later system the microprocessor will then be connected to SUMEX where it will become a slave to an AI program running on SUMEX. The SUMEX program will use a model appropriate for the organ which will form the basis of an initial hypothesis about the shape of the organ. This hypothesis will be refined at first by asking the user relevant clinical questions such as (for the fetus) the gestational age, the lie of the fetus in the abdomen and complicating medical factors. This kind of information is the same as that used by the clinician before he even places the scan head on the patient. The model will then be used to request those scans from the video disk which have the best chance of giving useful information. Heuristics based on the protocols used by clinicians during an exam will be incorporated since clinicians tend to collect scans in a manner which gives the most information about the organ. For each requested scan a two-dimensional tolerance region (or plan) derived from the model will be sent to the microcomputer. The requested scan will be retrieved from the video disk, digitized into a frame buffer, and the plan used to direct a border recognition process that will determine the organ outline on the scan. The resulting outline will be sent to SUMEX where it will be used to update the model. The scan requesting process will be continued until it is judged that enough information has been collected. The final model will then be used to determine volume and other quantitative parameters, and will be displayed in three dimensions.

We believe that this hypothesize verify method is similar to that used by clinicians when they perform an ultrasound exam. An initial model, based on clinical evidence and past experience, is present in the clinician's mind even before he begins the exam. During the exam this model is updated by collecting scans in a very specific manner which is known to provide the maximum amount of information. By building an ultrasound imaging system which closely resembles the way a physician thinks we hope to not only provide a useful diagnostic tool but also to explore very fundamental questions about the way people see.

We developed this system in phases, starting with an earlier version developed at the University of Washington. During the first phase the previous system was adapted and extended to run in the SUMEX environment. Clinical studies were done to determine its effectiveness in predicting fetal weight. In the second phase computer vision techniques were used to solve some of the problems observed in the clinical trials on the first phase.

B. Medical Relevance and Collaboration

This project was developed in collaboration with the Ultrasound Division of the Department of Obstetrics at Stanford, of which W.D. McCallum is the director.

Fetal weight is known to be a strong indicator of fetal well-being: small babies generally do more poorly than larger ones. In addition, the rate of growth is an important indicator: fetuses which are "small-for-dates" tend to have higher morbidity and mortality. It is thought that these small-for-dates fetuses may be suffering from placental insufficiency, so that if the diagnosis could be made soon enough early delivery might prevent some of the complications. In addition such growth curves would aid in understanding the normal physiology of the fetus. Several attempts have been made to use ultrasound for predicting fetal weight since ultrasound is painless, noninvasive, and apparently risk-free. These techniques generally use one or two measurements such as abdominal circumference or biparietal diameter in a multiple regression against weight. We previously studied several of these methods and concluded that the most accurate were about +/-200 gms/kg, which is not accurate enough for adequate growth curves (the fetus grows about 200 gms/week). The method we developed is based on the fact that fetal weight is directly related to volume since the density of fetal tissue is nearly constant. As part of this research we showed that by utilizing three dimensional information more accurate volumes and hence weights can be obtained.

In addition to fetal weight, the first implementation of this system was evaluated for its ability to determine other organ volumes in vitro. In collaboration with Dr. Richard Popp of the Stanford Division of Cardiology we evaluated the system on in vitro kidneys and latex molds of the human left ventricle. Left ventricular volumes are routinely obtained by means of cardiac catheterization in order to help characterize left ventricular function. Attempts to determine ventricular volume using one or two dimensional information from ultrasound has not demonstrated the accuracy of angiography. Therefore, three-dimensional information should provide a more accurate means of non-invasively assessing the state of the left ventricle.

C. Highlights of Research Progress

This section will summarize the major accomplishments of this project during its tenure on SUMEX. These accomplishments are described in detail in the Ph.D. dissertation of J. Brinkley, which is listed in the section on recent publications. The completion of the Ph.D. is the reason this project is now being discontinued.

The initial accomplishment was development of a microprocessor-based data acquisition system for acquiring a series of ultrasound images from a patient. The data acquisition system was designed to allow data to be acquired rapidly because the patient and organ must remain motionless while data is acquired. For this reason the exam was divided into 3 passes: patient exam, data entry and data analysis.

In the first pass, video ultrasound images are acquired from a commercial ultrasound scanner and stored on a videotape recorder, while position information from the locator is stored on floppy disk. In the data entry pass these scans are recalled from the tape recorder and outlined with the light pen. In the third pass the positions and outlines are sent to SUMEX, where the data analysis occurs. Software at SUMEX generates the 3D position of all outline points and allows them to be displayed graphically.

Before it was possible to use the data it was necessary to ascertain the accuracy of the 3D points. The accuracy of 3D point determination was found to be .6 cm. Individual sources of this error were analyzed and found to come about equally from the scanner resolution and the locator. These results were reported in Brinkley, Muramatsu et al., 1982.

The 3D points form the input to the modelling system. A regular mathematical model must be fitted to the arbitrary data in order to allow accurate volumes to be calculated. Two types of modelling system were developed: a "data-driven" system, which uses simple numerical techniques to interpolate a model to the data, and a "knowledge-driven" system which uses artificial intelligence techniques to overcome many of the deficiencies in the data-driven approach.

A detailed description and engineering evaluation of the data-driven approach can be found in Brinkley, Muramatsu et al., 1982. In the data driven system a series of regularly spaced scans are fitted to whatever data is present. The computer has no knowledge of what it is looking at. Engineering evaluations of this system were done on balloons, kidneys and molds of the human left ventricle, imaged in a water bath. For all three types of objects calculated volumes were generally within 5 percent of measured volume. These results provided justification for continued development, and showed more promise than standard clinical techniques which only use one or two measurements and an assumed shape.

The data-driven system was next evaluated for its ability to predict fetal weight, first *in vitro*, then *in utero*. The *in vitro* results are described in Brinkley, McCallum et al, 1982. In this study the relationship between measured weight and measured volume for a series of 26 dead neonates was shown to be highly linear, thus justifying the use of volume as a measure of fetal weight. The ability of volumes found by head and trunk reconstructions to predict fetal weight was then determined, and found to be quite good (R=.985).

The system was then used to predict fetal weight *in utero* as described in Brinkley, McCallum et al, 1983. Forty-one pregnant women were imaged within 48 hours of delivery. A total of 19 ultrasonic measurements were made, including head and trunk volume by reconstruction, as well as many simpler measurements utilized in the literature. These measurements were compared with weight measured at birth. The best combination of measurements was found to be a product of three head diameters, a product of three trunk diameters and trunk volume by reconstruction, giving a standard error of 69 g/kg (against natural log of birthweight). The most popular method in the literature gave a standard error of 106 g/kg suggesting that 3D information could improve weight prediction by about 30 percent.

However, further analysis showed that if the trunk volume by reconstruction was not included the standard error was still 73 g/kg, showing that the volumes by reconstruction were not that useful. This observation led to an evaluation of some of the problems in the data driven system, which in turn led to the need for an artificial intelligence approach.

The basic problems with the data-driven system were noise, missing data and awkwardness. Missing data was especially a problem in the term fetus since it was often impossible to visualize the fetal head and neck. If these data were not present the resulting volume would be too small because the computer had no way of knowing that it should interpolate an approximate neck or rump volume. The awkwardness came from the fact that it was necessary to outline all the scans with a light pen - usually about 90 minutes for a head and trunk reconstruction.

These problems were all related to the fact that the computer had no knowledge of what it was looking at. The goal of the knowledge-driven program was to give the computer the kind of anatomic knowledge that a radiologist utilizes in order to overcome deficiencies in the data.

The knowledge-driven system is described in Brinkley 1983 and Brinkley 1985. The system was implemented and tested on two shape classes of balloons (round and long-thin). For each balloon class a training set of similarly-shaped balloons was used to give the computer knowledge of the given shape. This training set consisted of ultrasonic reconstructions obtained by the previous system. The knowledge was then used to analyze ultrasound data from a similarly-shaped balloon which was not part of the training set. The initial input to the system consisted of the three-dimensional positions and orientations of a series of ultrasound slices. These slices were previously acquired manually and stored on a video tape recorder. The system was also given the two endpoints of the balloons, which allowed a reference coordinate system to be established. The balloon endpoints interacted with the shape knowledge to define an initial tolerance region, within which the system expected the actual balloon surface to be found. The system's best guess as to the location of the actual balloon surface was the middle of the tolerance region.

Once the initial tolerance region was established an hypothesize-verify paradigm was employed to alternately request a particular ultrasound slice, to provide a tolerance region for an edge detector on that slice, to manually acquire the border of the balloon on that slice, and to update the model by combining the new data with the shape knowledge. This process continued until it was judged that additional slices could contribute no new information.

For an example round balloon (measured volume 267 cc) the initial best guess volume after specifying the endpoints was 242 cc. After one slice best guess volume was 279 cc. After nine slices (out of a possible 30) the system judged that no more slices would be useful; best guess volume was 265 cc. For a different training set of long-thin balloons the final best guess volume for a new reconstruction, after 9 out of a possible 22 slices, was 459 cc, measured volume 461 cc. These results show that learned shape knowledge allowed the system to form a reasonable guess as to the location of the balloon surface even after only two endpoints had been specified.

The overall conclusions of this research are (1) three-dimensional ultrasound data provides accurate volumes at least *in vitro*, (2) 3D data may improve fetal weight prediction by approximately 30 percent (3) use of artificial intelligence techniques, when further developed, hold promise for greatly improving the performance of a three-dimensional organ modelling system.

D. Recent Publications

1. Brinkley, J.F., Muramatsu, S.K., McCallum, W.D. and Popp, R.L.: *In vitro evaluation of an ultrasonic three-dimensional imaging and volume system.* Ultrasonic Imaging, 4:126-139, 1982.

2. Brinkley, J.F., McCallum, W.D., Muramatsu, S.K. and Liu, D.Y.: *Fetal weight estimation from ultrasonic three-dimensional head and trunk reconstructions: Evaluation in vitro.* Amer. J. Obstet. Gynecol. 144(6):715-721, 1982.

3. Brinkley, J.F., McCallum, W.D., Muramatsu, S.K., and Liu, D.Y.: *Fetal weight estimation from lengths and volumes found by ultrasonic three-dimensional measurements.* J. Ultrasound Med. 3:163-168, 1983.

4. Brinkley, J.F.: *Learned shape knowledge in ultrasonic three-dimensional organ modelling.* Second place, student paper competition, Symposium on Computer Applications in Medical Care, Baltimore, October 23-26, 1983.

5. Brinkley, J.F.: *Ultrasonic three-dimensional organ modelling.* Ph.D. Dissertation, Stanford University, Stanford Computer Science Technical report STAN-CS-84-1001, 1984.

6. Brinkley, J.F.: *Knowledge-driven ultrasonic three-dimensional organ modelling.* To be published in *IEEE Trans. Pattern Analysis and Machine Intelligence*, Summer 1985.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

We collaborated more with medical people than anyone else. The project was located in the Obstetrics Department at Stanford where W.D. McCallum manages the ultrasound patients. We also collaborated with Dr. Richard Popp in the Division of Cardiology at Stanford.

B. Sharing and Interactions with SUMEX projects

Mostly personal contacts with the Heuristic Programming Project and Medical Information Science Program at Stanford. The message facilities of SUMEX have been especially useful for maintaining these contacts.

C. Critique of Resource Management

In general SUMEX has been a very usable system, and the staff has been very helpful.

III. RESEARCH PLANS

A. Project Goals and Plans

The major conclusion from the research leading to the Ph.D. is that the current hardware we use for three-dimensional location is not accurate enough to permit further work on organ modelling. For this reason I have proposed several alternative methods of utilizing 3D medical image data, including 3D CT, NMR or ultrasound. All these modalities produce 3D arrays of data which would be much easier to use than arbitrary slices.

Given this type of data, fairly straightforward extensions of the model representation developed for balloons could be used for the heart or kidney. The basic idea would be to have the human operator indicate three organ landmarks within the 3D data, then let the computer utilize learned shape knowledge to selectively "biopsy" portions of the 3D data in order to define the actual organ instance. Since the data would be available as a 3D array, the edge detection process could take place along a one-dimensional tolerance region rather than on a two-dimensional slice. Since all forms of medical images are becoming available as 3D arrays this seems like a better approach than the selection of individual slices.

Depending on the interest of engineers in providing 3D data much of the AI modelling could still be done on SUMEX. Many of the AI techniques could also be developed for 2D images for knowledge-driven border detection. However, there are no plans to continue this research at present.

B. Justification and requirements for continued SUMEX use

The goals of this project seem to be compatible with the general goals of SUMEX, i.e., to develop the uses of artificial intelligence in medicine. The problem of three-dimensional modelling is a very general one which is probably at the heart of our ability to see. By developing a medical imaging system that models the way clinicians approach a patient we should not only develop a useful clinical tool but also explore some very fundamental problems in AI.

The availability of a large well supported facility like SUMEX was very useful for developing this system.

C. Needs and plans for other computing resources beyond SUMEX-AIM

Judging from our present experience it appears that SUMEX could not handle the amount of data required for image processing on digitized ultrasound scans. The recent advent of relatively powerful microprocessors and personal LISP machines makes these machines very attractive for further development. SUMEX could still act as a communications crossroads, however.

D. Recommendations

Since any further research on this project would require dedicated image processors we would hope to see these kind of systems being developed by the SUMEX resource. Projects that would be of direct interest are networks (such as ETHERNET), personal computer stations, graphics displays, etc.

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 80 common benign and malignant diseases of the lymph node 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.

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 10 new cases of malignant lymphomas are diagnosed each year. Even in a university hospital, only approximately 100 new patients are diagnosed every year.

Because of the limited numbers of cases seen, pathologists may not be conversant with the differential diagnoses consistent with each of the histologic features of the lymph node; they may lack familiarity with the complete spectrum of the histologic findings associated with a wide range of diseases. In addition, pathologists may be unable to fully comprehend the conflicting concepts and terminology of the different classifications of non-Hodgkin's lymphomas, and may not be cognizant of the significance of the immunologic, cell kinetic, cytogenetic, and immunogenetic data associated with each of the subtypes of the non-Hodgkin's lymphomas.

In order to promote the accuracy of the knowledge base development we will have participants for multiple institutions collaborating on the project. Dr. Nathwani will be joined by experts from Stanford (Dr. Dorfman), St. Jude's Children's Research Center -- Memphis (Dr. Berard) and City of Hope (Dr. Burke).

C. Highlights of Research Progress

C.1 Accomplishments This Past Year

Since the project's inception in 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.

C.1 The PATHFINDER knowledge base

The basic building block of the PATHFINDER knowledge base is the disease profile or *frame*. The disease frame consists of *features* useful for diagnosis of lymph node diseases. Currently these features include histopathologic findings seen in both low- and high-power magnifications. Each feature is associated with a list of exhaustive and mutually exclusive *values*. For example, the feature *pseudofollicularity* can take on any one of the values *absent*, *slight*, *moderate*, or *prominent*. These lists of values give the program access to *severity* information. In addition, these lists eliminate obvious interdependencies among the values for a given feature. For example, if pseudofollicularity is *moderate*, it cannot also be *absent*.

Evoking strengths and frequencies are associated with each feature-value pair in a disease profile. We are experimenting with different scales for scoring each feature-value pair, and several methods for combining the scores to form a differential diagnosis. A disease-independent import is also assigned to each feature-value but only a two-valued scale is used. This is because, in PATHFINDER, imports are only used to make boolean or yes/no decisions (see below). In addition to import, PATHFINDER utilizes the concept of *classic* features for a disease -- within each disease frame, the pathologist marks those feature-value pairs which are considered to be part of the classic pattern of the disease.

The PATHFINDER knowledge base contains information about obvious association between features. This information is of the form: "Don't ask about feature x unless feature y has certain values." For example, it wouldn't make sense to ask about the degree or range of follicularity if there are no follicles in the tissue section. The feature links also serve to identify interdependencies among features. Feature interdependence is a problem because it can lead to inaccuracies in scoring hypotheses.

The prototype knowledge base was constructed by Dr. Nathwani. During the beginning part of 1984, we organized two meetings of the entire team 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.

D. Publications Since January 1984

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.

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 expert, 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, 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 HP9836 workstation. We will need additional file space on the 2060 as we quadruple the size of our knowledge base. We will continue to require access to the 2060 for communication purposes, access to other programs, and for file storage and archiving.

D. Recommendations for Future Community and Resource Development

We encourage the continued exploration by SUMEX of the interconnection of workstations within the mainframe computer setting. We will need to be able to quickly move a program from workstation to workstation, or from workstation back and forth to the mainframe. Software tools that would help the transfer of programs from one type of workstation to another would also be quite useful. Until the type of workstations that we are using in this research becomes inexpensive ($5000 or less), 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.
Manfred Kochen, 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 has been a growing emphasis in American psychiatry on careful diagnosis using clearly defined clinical criteria (Feighner, et al., 1972; Spitzer, et al., 1975, 1980; Feinberg and Carroll, 1982, 1983). These efforts have led to several sets of criteria for the diagnosis of psychiatric disorders. The "St. Louis" criteria (Feighner, et al., 1972) were succeeded by the Research Diagnostic Criteria (RDC), formulated by researchers from St. Louis and New York (Spitzer, et al., 1975). The RDC led directly to the criteria that are now quasi-official in American psychiatry, DSM-III (Spitzer, et al., 1980). All of these criteria lists were based on a combination of clinical opinion and literature review, and use a decision-tree approach to making a diagnosis. These diagnostic systems have been shown to be acceptably reliable, but their validity remains untested. Other groups have used a multivariate statistical approach to diagnosis. Roth and his colleagues (Carney, et al., 1965) published a discriminant index for distinguishing "endogenous" from "neurotic" depressed patients. This work was repeated by Kiloh, et al. (1972) with much the same results, confirming the findings of Carney, et al. (1965).

We have done similar work, deriving two discriminant indices for separating endogenous depressed patients (unipolar or bipolar) from nonendogenous (neurotic) patients. We cross-validated these indices in separate groups of patients, and also validated them against an external standard, the dexamethasone suppression test (Feinberg and Carroll, 1982, 1983). At the same time, we and others have been further developing this and other biological measures that may differentiate between patients with endogenous and nonendogenous depression. These include neuroendocrine tests such as the dexamethasone suppression test (DST) and quantitative studies of sleep using EEG. Carroll, et al. (1981) have shown that the DST is abnormal in about 67%

of patients with endogenous depression (melancholia) and only 5-10% with nonendogenous (neurotic) depression. Kupfer, et al. (1978) and Feinberg, et al. (1982) have similar results with EEG studies of sleep. These biological markers may be useful for routine clinical use, and can certainly be used as external validating criteria to test the performance of different clinical diagnostic methods, including those mentioned above. Furthermore, we have developed biological criteria for "definitely endogenous" depression and "definitely nonendogenous" depression based on DST and sleep EEG. (Carroll, et al., 1980). Our goal is to use these criteria as an external validating criterion for assessing the performance of various new or different diagnostic schemes, in particular an expert system of the sort we are developing.

C. Highlights of Research Progress

We examined two other SUMEX-based psychiatry projects, the BLUEBOX project of Mulsant and Servan-Schreiber (1984), and the HEADMED project of Heiser and Brooks (1978, 1980). Mulsant and Servan-Schreiber visited us at Michigan and discussed the rationale and progress of their project. Heiser also visited with us and agreed to collaborate with our project as a consultant.

At Michigan, we encoded the Hamilton Rating Scale (Hamilton, 1967) into EMYCIN rules. This is the standard scale (in English) for rating the severity of depression, and many of the items in it are relevant to our consultant program. We moved our work to the AGE system, breaking the Hamilton scale into its component subscales and adding other components to determine patient demographic information, personal and family psychiatric history, and other rating scale information. We then introduced other knowledge sources to construct a differential diagnosis list for psychiatric illnesses based on our expert's taxonomy and methods. We are now focussing on rules that discriminate endogenous from non-endogenous depression. Concurrently we are developing a treatment knowledge base on a LISP workstation. Thus far, the treatment knowledge base contains information about drug therapies, including types, dosages, activities, interactions, and side effects.

We have conducted interviews with patients recently admitted to the University of Michigan Adult Psychiatric Hospital. They are interviewed by Feinberg and the interviews are observed by Lindsay plus a group of psychiatric residents, psychiatrists and psychologists. After the interview, Feinberg is debriefed by Lindsay, and then the others discuss the case. These data are the initial source of the expert knowledge base for our consultant.

D. List of Relevant Publications

This project has not yet produced any publications. The following list contains the references cited above, including our previous publications relevant to the RxDx Project.

1. Carney, M. W. P., Roth, M. and Garside, R. F.:*The diagnosis of depressive syndromes and the prediction of ECT response*, Brit. J. Psychiatry, 111, 659-674, 1965.

2. Carroll, B. J., Feinberg, M., Greden, J. F., Haskett, R. F., James, N. McI., Steiner, M., and Tarika, J.: *Diagnosis of endogenous depression: Comparison of clinical, research, and neuroendocrine criteria*, J. Affect Dis., 2, 177-194, 1980.

3. Carroll, B. J., Feinberg, M., Greden, J. F., Tarika, J., Albala, A. A., Haskett, R. F., James, N. McI., Kronfol, Z., Lohr, N., Steiner, M., de Vigne, J-P, and Young, E.:*A specific laboratory test for the diagnosis of melancholia*, Standardization, validation, and clinical utility. Arch. Gen. Psychiatry, 38, 15-22, 1981.

4. Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G., and Munoz, R.: *Diagnostic criteria for use in psychiatric research*, Arch. Gen. Psychiatry, 26, 57-63, 1972.

5. Feinberg, M. and Lindsay, R. K.: *Expert systems*. Proceedings of the NCDEU Annual Meeting, Key Biscayne, Florida, May 1985.

6. Feinberg, M. and Carroll, B. J.: *Separation of subtypes of depression using discriminant analysis: I. Separation of unipolar endogenous depression from non-endogenous depression*, Brit. J. Psychiatry, 140, 384-391, 1982.

7. Feinberg, M. and Carroll, B. J.:*Separation of subtypes of depression using discriminant analysis. II. Separation of bipolar endogenous depression from nonendogenous ("neurotic") depression*, J. Affective Disorders, 5, 129-139, 1983.

8. Feinberg, M. and Carroll, B.J.: *Biological markers for endogenous depression in series and parallel*, Biological Psychiatry 19:3-11, 1984.

9. Feinberg, M. and Carroll, B.J.: *Biological and nonbiological depression*, Presented at Annual Meeting of the Society of Biological Psychiatry, Los Angeles, May, 1984, Abstract #81.

10. Feinberg, M., Gillin, J. C., Carroll, B. J., Greden, J. F., and Zis, A. P.:*EEG studies of sleep in the diagnosis of depression*, Biological Psychiatry, 17, 305-316, 1982.

11. Heiser, J. F. and Brooks, R. E.:*Design considerations for a clinical psychopharmacology advisor*, Proc. Second Annual Symp. on Computer Applications in Medical Care. New York: IEEE, 1978, 278-285.

12. Heiser, J. F. and Brooks, R. E.:*Some experience with transferring the MYCIN system to a new domain*, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-2, No. 5, 477-478, 1980.

13. Kiloh, L. G., Andrews, G., and Neilson, M.:*The relationship of the syndromes called endogenous and neurotic depression*, Brit. J. Psychiatry, 121, 183-196, 1972.

14. Kupfer, D. J., Foster, F. G., Coble, P., McPartland, R. J., and Ulrich, R. F.:*The application of EEG sleep for the differential diagnosis of affective disorders*, Am. J. Psychiatry, 135, 69-74, 1978.

15. Mulsant, B. and Servan-Schreiber, D.:*Knowledge engineering: A daily activity on a hospital ward*, Computers in Biomedical Research, 1984.

16. Spitzer, R. L., Endicott, J. and Robins, E.: *Research diagnostic criteria*, (2d ed.) New York State Department of Mental Hygiene, New York Psychiatric Institute, Biometrics Research Division, 1975.

17. Spitzer, R. L.: (Ed.)*Diagnostic and statistical manual of mental disorders*, (3d ed.). Washington, D. C.: American Psychiatric Association, 1980.

18. Van Melle, W.:*The EMYCIN Manual*, Computer Science Department, Stanford University, Report HPP-81-16, 1981.

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

Our use of EMYCIN and AGE has been of major importance. In addition, we have worked with Dr. Larry Fagan to learn about his Pathfinder program. We used that program, on SUMEX, to obtain some information for the RxDx project by applying it to data we previously collected on depression symptom frequencies.

C. Critique of Resource Management

We have been using EMYCIN and AGE in our work, and have found these programs very valuable, saving us many hours of programming in LISP. There are some problems with them, many of which center around discrepancies between the versions described in the manuals and the versions actually running on SUMEX. We would suggest that software be more strongly supported than is now the case, if it and SUMEX are to be even more useful to beginners in AI in Medicine.

SUMEX itself has been invaluable. We don't have ready access to any other machine of equal computing power which also has a strongly supported LISP available. Specifically, the LISP compiler available on the Amdahl 5860 here differs from those used at major AI centers such as Stanford and MIT. We have also made good use of the ARPANET connections that SUMEX offers. Feinberg spent a month of his sabbatical working with Prof. Peter Szolovits at MIT, learning about AI in Medicine. This visit was arranged using computer mail through SUMEX. Lindsay and Feinberg were able to continue their collaborative work while the latter was in Cambridge, using the same medium. The alternative would have been days lost in the mails and many dollars spent on phone calls. We have also been able to get help with problems that arise with EMYCIN and AGE using computer mail.

Most of the limitations of SUMEX, and they are often severe, derive from the necessity to access it via TYMNET. Response time is often impossibly slow, and even at its best the delays are annoying and frustrating, even for editing and debugging. For example, editing is limited to a primitive line editor, since EMACS interacts with the network XON/XOFF handshaking in a disastrous way. The staff has not been helpful in solving these network related problems, probably because they do not have to live with them in their own interactions with the system. In any case, many of the problems are beyond the reach of the Sumex staff. The future of long-haul network collaborations depends critically on increased bandwidth and faster response times.

It would have been helpful to us to obtain the AGE system that runs on a Xerox 1108. However, the $530 price, though perhaps modest in comparison to its development costs, was beyond the reach of our budget. It would be helpful if distribution costs for software could be held under $100.

III. RESEARCH PLAN

A. Project Goals and Plans

Our immediate objective is to develop an expert system 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 MD'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 plan to move our system to local workstations as soon as we are able, 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 level for the next year.

C. Needs and Plans for Other Computing Resources

As our project evolves and we run into the limitations of the time-shared SUMEX facility, we anticipate employing different expert systems software. At this time, we are not at a stage to say exactly what that will be, but our project is not sufficiently large that we will be able to mount such a software development project ourselves, so we will depend on development and support elsewhere. Ultimately, when our consultant is made available for field trials and clinical use, it will need to be transported to a personal computer that is large enough to support the system yet inexpensive enough to be widely available. A LISP machine is an obvious candidate. While current prices of the necessary hardware are too high, computer prices are continuing to drop. Our design strategy is to avoid limiting ourselves and our aspirations to that which is affordable today; instead we will attempt to project the growth of our project and the price-performance curve of computing such that they meet at some reasonable point in the future.

D. Recommendations for Future Community and Resource Development

Valuable as the present SUMEX facilities are to us, they are in many ways limited and awkward to use. The 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 beyond the next year or eighteen months.

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.

Appendix A

Stanford Knowledge Systems Laboratory

ARTIFICIAL INTELLIGENCE RESEARCH IN THE KNOWLEDGE SYSTEMS LABORATORY (Incorporating the Heuristic Programming Project)

Stanford University
Department of Computer Science/Department of Medicine
April 1985

The Knowledge Systems Laboratory (KSL) is an artificial intelligence research laboratory of about 90 people -- faculty, staff, and students -- within the Departments of Computer Science and Medicine at Stanford University. KSL is the new name for the interdisciplinary AI research community that has evolved over the past two decades. Begun as the DENDRAL Project in 1965 and known as the Heuristic Programming Project from 1972 to 1984, the new organization reflects the increasing complexity and diversity of the research now under way. The KSL is a modular laboratory, consisting of five collaborating yet distinct groups with different research themes:

- **The Heuristic Programming Project** (HPP), Professor Edward A. Feigenbaum, scientific director -- blackboard systems, concurrent system architectures for AI, and the modeling of discovery processes. Executive director: Robert Engelmore. Research scientists: Harold Brown, Byron Davies, Bruce Delagi, Peter Friedland, Barbara Hayes-Roth, and H. Penny Nii. Consulting professor: Richard Gabriel.
- **The HELIX Group**, Professor Bruce G. Buchanan, scientific director -- machine learning, transfer of expertise, and problem solving. Faculty: Paul S. Rosenbloom (joint appointment, Computer Science and Psychology). Research scientists: James Brinkley, William J. Clancey, Barbara Hayes-Roth.
- **The Medical Computer Science (MCS) Group**, Professor Edward H. Shortliffe, scientific director (Department of Medicine with courtesy appointment in Computer Science) -- research on and advanced application of AI to medical problems; includes the Medical Information Sciences (MIS) program. Research scientist: Lawrence M. Fagan.
- **The Logic Group**, Professor Michael R. Genesereth, scientific director -- formal reasoning and introspective systems. Research scientist: Matthew L. Ginsberg.
- **The Symbolic Systems Resources Group** (SSRG), Thomas C. Rindfleisch, scientific director (joint appointment, Computer Science and Medicine) -- research on and operation of computing resources for AI research, including the SUMEX facility. Assistant director: William J. Yeager.

Tom Rindfleisch serves as KSL project director.

This brochure summarizes the goals and methodology of the KSL, its research and academic programs, its achievements, and the research environment of the laboratory.

Basic Research Goals and Methodology

Throughout a 20-year history, the KSL and its predecessors, DENDRAL and HPP, have concentrated on research in expert systems -- that is, systems using symbolic reasoning and problem-solving processes that are based on extensive domain-specific knowledge. The KSL's approach has been to focus on applications that are themselves significant real-world problems, in domains such as science, medicine, engineering, and education, and that also expose key, underlying AI research issues. For the KSL, AI is largely an empirical science. Research problems are explored, not by examining strictly theoretical questions, but by designing, building, and experimenting with programs that serve to test underlying theories.

The basic research issues at the core of the KSL's interdisciplinary approach center on the computer representation and use of large amounts of domain-specific knowledge, both factual and heuristic (or judgmental). These questions have guided our work since the 1960s and are now of central importance in all of AI research:

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?
2. **Knowledge acquisition.** How is knowledge acquired most efficiently -- whether from human experts, from observed data, from experience, or by discovery? How can a program discover inconsistency and incompleteness in its knowledge base? How can knowledge be added without perturbing the established knowledge base?
3. **Use of knowledge.** By what inference methods can many sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? How can knowledge be used intelligently, especially in systems with large knowledge bases, so that it is applied in an appropriate manner at the appropriate time?
4. **Explanation and tutoring.** How can the knowledge base and the line of reasoning used in solving a particular problem be explained to users? What constitutes a sufficient or an acceptable explanation for different classes of users? How can problem-solving systems be combined with pedagogical and user knowledge to implement intelligent tutoring systems?
5. **System tools and architectures.** What kinds of software tools and system architectures can be constructed to make it easier to implement expert programs with greater complexity and higher performance? What kinds of systems can serve as vehicles for the cumulation of knowledge of the field for the researchers?

Research and Academic Programs

CURRENT RESEARCH PROJECTS

The following list of projects now under way within the five KSL research groups gives a brief summary of the major goals of each project and lists the personnel (staff and Ph.D. candidates) directly involved. More complete information on individual projects can be obtained from the person indicated as the project contact. Inquiries should be addressed in care of:

Knowledge Systems Laboratory
Department of Computer Science
Stanford University
701 Welch Road, Building C
Palo Alto, CA 94304
415-497-3444

The Heuristic Programming Project

- **Advanced Architectures Project** -- Design a new generation of computer architectures to exploit concurrency in blackboard-based signal understanding systems.
  Personnel: Edward A. Feigenbaum (contact), Harold Brown, Byron Davies (TI), Bruce Delagi (DEC), Richard Gabriel, Penny Nii, Sayuri Nishimura, Jim Rice, Eric Schoen, Jerry Yan.
- **Knowledge-Based VLSI Design Project** -- Study the hierarchical design process involved in the development of complex very large scale integrated circuits.
  Personnel: Harold Brown (contact), Jerry Yan.
- **Blackboard Architecture Project** -- Integrate current knowledge about blackboard framework problem-solving systems and develop a domain-independent model that includes knowledge-based control processes.
  Personnel: Barbara Hayes-Roth (contact).
- MOLGEN -- Study the processes of scientific theory formation and modification, using recently developed models of genetic regulation as an example.
  Personnel: Peter Friedland (contact), Charles Yanofsky (Biological Science), Peter Karp.

The HELIX Group

- PROTEAN -- Study complex symbolic constraint-satisfaction problems in the blackboard framework with application to protein structure determination from nuclear magnetic resonance data.
  Personnel: Bruce Buchanan (contact), Oleg Jardetzky (Nuclear Magnetic Laboratory), Jim Brinkley, Barbara Hayes-Roth, Russ Altman, Olivier Lichtarge.
- NEOMYCIN/GUIDON2 -- Develop knowledge representation and explanation capabilities for the computer-aided teaching of diagnostic reasoning.
  Personnel: Bill Clancey (contact), Stephen Barnhouse, Diane Hasling, David C. Wilkins.
- SOAR -- Develop a general production-system-based problem-solving architecture that integrates reasoning, domain expertise, learning, and planning of problem-solving strategies.
  Personnel: Paul Rosenbloom (contact), Andrew Golding, Amy Unruh.
- **Knowledge Acquisition Studies** -- Study the processes for transferring knowledge into a computer program, including learning by induction, analogy, watching, chunking, reading, and discovery.
  Personnel: Bruce Buchanan (contact), Li-Min Fu, Russell Greiner, Ramsey Haddad, David C. Wilkins.

The Medical Computer Science Group

- ONCOCIN -- Develop knowledge-based systems for the administration of complex medical treatment protocols such as those encountered in cancer chemotherapy. *Personnel*: Ted Shortliffe (contact), Charlotte Jacobs (Oncology), Larry Fagan, David Combs, Gregory Cooper, Jay Ferguson, Christopher Lane, Janice Rohn, Homer Chin, Holly Jimison, Curt Langlotz, Mark Musen, Glenn Rennels.

- PATHFINDER -- Develop a knowledge-based system for diagnosis of lymph node pathology.
  Personnel: Ted Shortliffe, Bharat Nathwani (USC), Larry Fagan (contact), David Heckerman, Eric Horvitz.

The Logic Group

- **Metalevel Representation System (MRS)** -- Study logic-based introspective programs that can reason about and control their own problem-solving activities.
  Personnel: Mike Genesereth (contact), Matt Ginsberg, Russ Greiner, Ben Grosof, Yung-Jen Hsu, David E. Smith, Devika Subramanian, Richard Treitel.
- **The DART/HELIOS Project** -- Study an integrated design environment that includes capabilities for design specification, refinement, and validation; fabrication engineering; and failure diagnosis and testing.
  Personnel: Mike Genesereth (contact), Glenn Kramer (Fairchild), Narinder Singh.
- **Intelligent Agent Project** -- Study planning and problem-solving activities for an intelligent interface between human users and complex computing environments.
  Personnel: Mike Genesereth (contact), Matt Ginsberg, Jeff Finger, Jeff Rosenschein, Jock Mackinlay, Vineet Singh.
- **Intelligent Task Automation** -- Build a program that can use the description of a manufacturing task to develop a plan by which a robot can carry out the task.
  Personnel: Mike Genesereth (contact), Matt Ginsberg, Jeff Finger, David E. Smith, Richard Treitel.

The Symbolic Systems Resources Group (SSRG)

- **SUMEX-AIM Resource** -- Develop and operate a national computing resource for biomedical applications of artificial intelligence in medicine and for basic research in AI at KSL.
  Personnel: Tom Rindfleisch (contact), Bill Croft, Frank Gilmurray, Christopher Schmidt, Andrew Sweer, Israel Torres, Bob Tucker, Nicholas Veizades, Bill Yeager.
- **Financial Resource Management** -- Develop an expert system for financial resource planning.
  Personnel: Tom Rindfleisch (contact), Bruce Buchanan.

Other Projects

The KSL also has close ties to collaborative projects. These include PIXIE, developing an intelligent tutoring system, under Derek Sleeman in the School of Education, and RADIX, studying discovery of knowledge from databases, under Bob Blum in Computer Science.

STUDENTS AND SPECIAL DEGREE PROGRAMS

Graduate students are an essential part of the research productivity of the KSL. Currently 41 students are working with our projects centered in Computer Science and another 12 students are working with the MCS/MIS programs in Medicine. Of the 41 working in Computer Science, 25 are working toward Ph.D. degrees, and 16 are working toward M.S. degrees. A number of students are pursuing interdisciplinary programs and come from the Departments of Engineering, Mathematics, Education, and Medicine.

Because of the highly interdisciplinary and experimental nature of KSL research, two special degree programs have been established:

The **Medical Information Sciences** (MIS) program is an interdepartmental program approved by Stanford University in 1982. It offers instruction and research opportunities leading to the M.S. or Ph.D. degree in medical information sciences, with an emphasis on either medical computer science or medical decision science. The program, directed by Ted Shortliffe and co-directed by Larry Fagan, is formally administered by the School of Medicine, but the curriculum and degree requirements are coordinated with the Dean of Graduate Studies and the Graduate Studies Committee of the University. The program reflects our local interest in the interconnections between computer science, artificial intelligence, and medical problems. Emphasis is placed on providing trainees with a broad conceptual overview of the field and with an ability to create new theoretical and practical innovations of clinical relevance.

The **Master of Science in Computer Science: Artificial Intelligence** (MS:AI) program is a terminal professional degree offered for students who wish to develop a competence in the design of substantial knowledge-based AI applications but who do not intend to obtain a Ph.D. degree. The MS:AI program is administered by the Committee for Applied Artificial Intelligence, composed of faculty and research staff of the Computer Science Department. Normally, students spend two years in the program with their time divided equally between course work and research. In the first year, the emphasis is on acquiring fundamental concepts and tools through course work and and project involvement. During the second year, students implement and document a substantial AI application project.

Academic and Research Achievements

The primary products of our research are scientific publications on the basic research issues that motivate our work, computer software in the form of the expert systems and AI architectures we develop, and the students we graduate who continue AI research in other academic and industrial laboratories.

The KSL has averaged publishing more than 45 research papers per year in the AI literature, including journal articles, theses, proceedings articles, and working papers. In addition, many talks and invited lectures are given annually. In the past few years, 11 major books have been published by KSL faculty, staff, and former students, and several more are in progress. Those recently published include:

- *Heuristic Reasoning about Uncertainty: An AI Approach*, Cohen, Pitman, 1985.
- *Readings in Medical Artificial Intelligence: The First Decade*, Clancey and Shortliffe, Addison-Wesley, 1984.
- *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Buchanan and Shortliffe, Addison-Wesley, 1984.
- *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*, Feigenbaum and McCorduck, Addison-Wesley, 1983.
- *Building Expert Systems*, F. Hayes-Roth, Waterman, and Lenat, eds., Addison-Wesley, 1983.

- *System Aids in Constructing Consultation Programs: EMYCIN*, van Melle, UMI Research Press, 1982.
- *Knowledge-Based Systems in Artificial Intelligence: AM and TEIRESIAS*, Davis and Lenat, McGraw-Hill, 1982.
- *The Handbook of Artificial Intelligence*, Volume I, Barr and Feigenbaum, eds., 1981; Volume II, Barr and Feigenbaum, eds., 1982; Volume III, Cohen and Feigenbaum, eds., 1982; Kaufmann.
- *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project*, Lindsay, Buchanan, Feigenbaum, and Lederberg, McGraw-Hill, 1980.

Our laboratory has pioneered in the development and application of AI methods to produce high-performance knowledge-based programs. Programs have been developed in such diverse fields as analytical chemistry (DENDRAL), infectious disease diagnosis (MYCIN), cancer chemotherapy management (ONCOCIN), pulmonary function evaluation (PUFF), machine fault diagnosis (DART), VLSI design (KBVLSI/PALLADIO), and molecular biology (MOLGEN). Some of these programs rival human experts in solving problems in restricted domains. A number of projects have developed generalized software tools for representing and using knowledge; of these, EMYCIN, AGE, MRS, and BB1 are available to outside research groups. Some of our systems and tools (e.g., DENDRAL, PUFF, UNITS, and EMYCIN) are now also being adapted for commercial development and use in the burgeoning AI industry.

Following our lead in work on biomedical applications of AI and the development of the SUMEX-AIM computing resource, a nationally recognized community of academic projects on AI in medicine has grown up.

Central to all KSL research are our faculty, staff, and students. These people have been recognized internationally for the quality of their work and for their continuing contributions to the field. KSL members participate extensively in professional organizations, government advisory committees, and journal editorial boards. They have held major managerial posts and conference chairmanships in both the American Association for Artificial Intelligence (AAAI) and the International Joint Conference on Artificial Intelligence (IJCAI).

Several KSL faculty and former students have received significant honors. In 1976, Ted Shortliffe received the Association of Computing Machinery Grace Murray Hopper award. In 1977, Doug Lenat received the IJCAI Computers and Thought award, and in 1978, Ed Feigenbaum received the National Computer Conference Most Outstanding Technical Contribution award. In 1981, Ted Shortliffe's book *Computer-Based Medical Consultation: MYCIN* was identified as the most frequently cited work in the IJCAI-81 proceedings. In 1982, Doug Lenat won the Tioga prize for the best AAAI conference paper while Mike Genesereth received honorable mention. In 1983, Ted Shortliffe was named a Kaiser Foundation faculty scholar, and Tom Mitchell received the IJCAI Computers and Thought award. In 1984, Randy Davis and Doug Lenat were named among the 100 most promising U.S. scientists under 40 by a prestigious scientific panel assembled by Science Digest. Also in 1984, Ed Feigenbaum was elected a fellow of the American Association for the Advancement of Science (AAAS), and he and Ted Shortliffe were elected fellows of the American College of Medical Informatics.

KSL Research Environment

Funding -- The KSL is supported solely by sponsored research and gift funds. We have had funding from many sources, including DARPA, NIH/NLM, ONR, NSF, NASA, and foundations and industry. Of these, DARPA and NIH have been the most substantial and long-standing sources of support. All, however, have made complementary contributions to establishing an effective overall research environment that fosters interchanges at the intellectual and software levels and that provides the necessary physical computing resources for our work.

Computing Resources -- Under the Symbolic Systems Resources Group, the KSL develops and operates its own computing resources tailored to the needs of its individual research projects. Current computing resources are a networked mixture of mainframe host computers, Lisp workstations, and network utility servers, reflecting the evolving hardware technology available for AI research. Our host machines include a DEC 2060 and 2020 running TOPS-20 (these are the core of the national SUMEX biomedical computing resource) and a VAX 11/780 running UNIX. Our growing complement of Lisp machines includes more than 25 Xerox 1100's, a Xerox Dorado, a Symbolics LM-2, eight Symbolics 3600's, and five Hewlett-Packard 9836's. Network printing, file, gateway, and terminal interface services are provided by dedicated machines ranging from VAX 11/750's to microprocessor systems. These facilities are integrated with other computer science resources at Stanford through an extensive Ethernet and to external resources through the ARPANET and Tymnet. Funding for these resources comes principally from DARPA and NIH.

Appendix B
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 Computer Science, TC135
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

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) 497-4879

KULIKOWSKI, Casimir, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-2006

LEDERBERG, Joshua, Ph.D.
President
The Rockefeller University
1230 York Avenue
New York, New York 10021
(212) 570-8080, 570-8000

LINDBERG, Donald A.B., M.D. (Past Adv Grp Chrmn)
Director, National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 02114
(617) 726-8311

MYERS, Jack D., M.D.
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 624-2649

AIM Advisory Group:

MYERS, Jack D., M.D. (Chairman)
School of Medicine
Scaife Hall, 1291
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
(412) 624-2649

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
Bethesda, Maryland 20205

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) 497-4879

KULIKOWSKI, Casimir, Ph.D.
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-2006

LEDERBERG, Joshua, Ph.D.
President
The Rockefeller University
1230 York Avenue
New York, New York 10021
(212) 570-8080, 570-8000

LINDBERG, Donald A.B., M.D.
Director, National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 02114
(617) 726-8311

MINSKY, Marvin, Ph.D.
Artificial Intelligence Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139
(617) 253-5864

MOHLER, William C., M.D.
Associate Director
Division of Computer Research and Technology
National Institutes of Health
Building 12A, Room 3033
9000 Rockville Pike
Bethesda, Maryland 20205
(301) 496-1168

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 Computer Science, TC135
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

SIMON, Herbert A., Ph.D.
Department of Psychology
Baker Hall, 339
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213
(412) 578-2787, 578-2000

Stanford Community Advisory Committee:

FEIGENBAUM, Edward A., Ph.D. (Chairman)
Heuristic Programming Project
Department of Computer Science
Margaret Jacks Hall
Stanford University
Stanford, California 94305
(415) 497-4879

LEVINTHAL, Elliott C., Ph.D.
Departments of Mechanical and Electrical Engineering
Building 530
Stanford University
Stanford, California 94305
(415) 497-9037

SHORTLIFFE, Edward H., M.D., Ph.D.
Principal Investigator - SUMEX
Medical Computer Science, TC135
Stanford University Medical Center
Stanford, California 94305
(415) 497-6970

Appendix C

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) 497-1997 (CLANCEY@SUMEX-AIM)

Bruce G. Buchanan, Ph.D.
Computer Science Department
701 Welch Road
Stanford University
Palo Alto, California 94304
(415) 497-0935 (BUCHANAN@SUMEX-AIM)

SOFTWARE AVAILABLE ON SUMEX

GUIDON--A system developed for intelligent computer-aided instruction. Although it was developed in the context of MYCIN's infectious disease knowledge base, the tutorial rules will operate upon any EMYCIN knowledge base.

NEOMYCIN--A consulation system derived from MYCIN, with the knowledge base greatly extended and reconfigured for use in teaching. In contrast with MYCIN, diagnostic procedures, common sense facts, and disease hierarchies are factored out of the basic finding/disease associations. The diagnostic procedures are abstract (not specific to any problem domain) and model human reasoning, unlike the exhaustive, top-down approach implicit in MYCIN's medical rules. This knowledge base will be used in the GUIDON2 family of instructional programs, being developed on D-machines.

REFERENCES

Clancey, W.J.: *Overview of GUIDON*. In A. Barr and E.A. Feigenbaum (Eds.), THE HANDBOOK OF ARTIFICIAL INTELLIGENCE, Vol. 2. William Kaufmann Assoc., Los Altos, CA, 1982. (Also to appear in J. of Computer-based Instruction)

Clancey, W.J.: *Methodology for building an intelligent tutoring system.* In Kintsch, Polson, and Miller, (Eds.), METHODS AND TACTICS IN COGNITIVE SCIENCE. L. Erlbaum Assoc., Hillsdale, NJ. 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) 497-2413

Contact: Dr. Peter FRIEDLAND@SUMEX-AIM
(415) 497-3728

The goal of the MOLGEN Project is to apply the techniques of artificial intelligence to the domain of molecular biology with the aim of providing assistance to the experimental scientist. Previous work has focused on the task of experiment design. Two major approaches to this problem have been explored, one which instantiates abstracted experimental strategies with specific laboratory tools, and one which creates plans in toto, heavily influenced by the role played by interactions between plan steps. As part of the effort to build an experiment design system, a knowledge representation and acquisition package--the UNITS System, has been constructed. A large knowledge base, containing information about nucleic acid structures, laboratory techniques, and experiment-design strategies, has been developed using this tool. Smaller systems, such as programs which analyze primary sequence data for homologies and symmetries, have been built when needed.

New work has begun on scientific theory formation, modification, and testing. This work will be done within the domain of regulatory genetics. We plan to explore fundamental issues in machine learning and discovery, as well as construct systems that will assist the laboratory scientist in accomplishing his intellectual goals.

SOFTWARE AVAILABLE ON SUMEX

SPEX system for experiment design.
UNITS system for knowledge representation and acquisition.
SEQ system for nucleotide sequence analysis.

REFERENCES

Friedland, P.E.: *Knowledge-based experiment design in molecular genetics*, (Ph.D. thesis). Stanford Computer Science Report, STAN-CS-79-771.

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

Stefik, M.J.: *An examination of a frame-structured representation system,*
Proc. Sixth IJCAI, Tokyo, August, 1979, pp. 845-852.

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

Stanford Project: ONCOCIN -- KNOWLEDGE ENGINEERING FOR ONCOLOGY CHEMOTHERAPY CONSULTATION

Principal Investigator: Edward H. Shortliffe, M.D., Ph.D.
Departments of Medicine and Computer Science
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (SHORTLIFFE@SUMEX-AIM)

Project Director: Dr. Lawrence M. Fagan

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 work is underway to transfer the program to professional workstations (rather than the central SUMEX computer) so that it can be implemented and evaluated at sites away from the University.

SOFTWARE AVAILABLE ON SUMEX

MYCIN-- A consultation system designed to assist physicians with the selection of antimicrobial therapy for severe infections. It has achieved expert level performance in formal evaluations of its ability to select therapy for bacteremia and meningitis. Although MYCIN is no longer the subject of an active research program, the system continues to be available on SUMEX for demonstration purposes and as a testing environment for other research projects.

EMYCIN-- The "essential MYCIN" system is a generalization of the MYCIN knowledge representation and control structure. It is designed to facilitate the development of new expert consultation systems for both clinical and non-medical domains.

ONCOCIN-- This system is in clinical use but is designed for special high speed terminals and therefore cannot be tested or demonstrated via network connections. Much of the knowledge in the domain of cancer chemotherapy is already well-specified in protocol documents, but expert judgments also need to be understood and modeled.

REFERENCES

Shortliffe, E.H., Scott, A.C., Bischoff, M.B., Campbell, A.B., van Melle, W. and Jacobs, C.D.: *ONCOCIN: An expert system for oncology protocol management*. Proc. Seventh IJCAI, pp. 876-881, Vancouver, B.C., August, 1981.

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

Langlotz, C.P. and Shortliffe, E.H.: *Adapting a consultation system to critique user plans*. Int. J. Man-Machine Studies 19:479-496, 1983.

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.

Stanford Project: PROTEAN Project

Principal Investigators: Oleg Jardetzky (JARDETZKY@SUMEX-AIM)
Nuclear Magnetic Resonance Lab, School of Medicine
Stanford University Medical Center
Stanford, California 94305

Bruce Buchanan, Ph.D. (BUCHANAN@SUMEX-AIM)
Computer Science Department
Stanford University
Stanford, California 94305

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

REFERENCES

1. Erman, L.D., Hayes-Roth, B., Lesser, V.R., Reddy, D.R.:*The HEARSAY-II Speech Understanding System: Integrating Knowledge to Resolve Uncertainty*. ACM Computing Surveys 12(2):213-254, June, 1980.

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

3. 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.

4. Hayes-Roth, B.:*A Blackboard Architecture for Control*. Artificial Intelligence In Press, 1985.

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

6. 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.

Stanford Project: RADIX -- DERIVING KNOWLEDGE FROM TIME-ORIENTED CLINICAL DATABASES

Principal Investigators: Robert L. Blum, M.D.
Departments of Medicine
and Computer Science
Stanford University
Stanford, California 94305
(415) 497-9421 (BLUM@SUMEX-AIM)

Gio C.M. Wiederhold, Ph.D.
Department of Computer Science
Stanford University
Stanford, California 94305
(415) 497-0685 (WIEDERHOLD@SUMEX-AIM)

The objective of clinical database (DB) systems is to derive medical knowledge from the stored patient observations. However, the process of reliably deriving causal relationships has proven to be quite difficult because of the complexity of disease states and time relationships, strong sources of bias, and problems of missing and outlying data.

The goal of the RADIX Project is to explore the usefulness of knowledge-based computational techniques in solving this problem of accurate knowledge inference from non-randomized, non-protocol patient records. Central to RADIX is a knowledge base (KB) of medicine and statistics, organized as a taxonomic tree consisting of frames with attached data and procedures. The KB is used to retrieve time-intervals of interest from the DB and to assist with the statistical analysis. Derived knowledge is incorporated automatically into the KB. The American Rheumatism Association DB containing records of 1700 patients is used.

SOFTWARE AVAILABLE ON SUMEX

RADIX--(excluding the knowledge base and clinical database) consists of approximately 400 INTERLISP functions. The following groups of functions may be of interest apart from the RADIX environment:

SPSS Interface Package -- Functions which create SPSS source decks and read SPSS listings from within INTERLISP.

Statistical Tests in INTERLISP -- Translations of the Piezer-Pratt approximations for the T,F, and Chi-square tests into LISP.

Time-Oriented Data Base and Graphics Package -- Autonomous package for maintaining a time-oriented database and displaying labelled time-intervals.

REFERENCES

Monograph

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.: *Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical database: The RX Project.* Computers and Biomedical Research 15(2):164-187, April, 1982.

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

Conference Proceeding

Blum, R.L.: *Modeling and encoding clinical causal relationships.* Proc. SCAMC83, IEEE, Baltimore, MD, October, 1983.

National AIM Project: CADUCEUS (formerly INTERNIST)

Principal Investigators: Jack D. Myers, M.D. (MYERS@SUMEX-AIM)
Harry E. Pople, Ph.D. (POPLE@SUMEX-AIM)
University of Pittsburgh
Pittsburgh, Pennsylvania 15261
Dr. Pople: (412) 624-3490
Dr. Myers: (412) 624-2649

The major goal of the CADUCEUS Project is to produce a reliable and adequately complete diagnostic consultative program in the field of internal medicine. Although this program is intended primarily to aid skilled internists in complicated medical problems, the program may have spin-off as a diagnostic and triage aid to physicians' assistants, rural health clinics, military medicine and space travel. In the design of CADUCEUS and its predecessor INTERNIST I, we have attempted to model the creative, problem-formulation aspect of the clinical reasoning process. The program employs a novel heuristic procedure that composes differential diagnoses, dynamically, on the basis of clinical evidence. During the course of a CADUCEUS or INTERNIST-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.

SOFTWARE AVAILABLE ON SUMEX

Versions of INTERNIST are available for experimental use, but the project continues to be oriented primarily towards research and development; hence, a stable production version of the system is not yet available for general use.

National AIM Project: CLIPR -- HIERARCHICAL MODELS OF HUMAN COGNITION

Principal Investigators: Walter Kintsch, Ph.D. (KINTSCH@SUMEX-AIM)
Peter G. Polson, Ph.D. (POLSON@SUMEX-AIM)
Computer Laboratory for Instruction in Psychological Research (CLIPR)
Department of Psychology
University of Colorado
Boulder, Colorado 80302
(303) 492-6991
Contact: Dr. Peter G. Polson (Polson@SUMEX-AIM)

The CLIPR Project is concerned with the modeling of complex psychological processes. It is comprised of two research groups. The prose comprehension group has completed a project that carries out the 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.

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

Fletcher, R. C. Understanding and solving word arithmetic problems: A computer simulation. Technical Report N0. 135, Institute of Cognitive Science, Colorado, 1984.

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

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

Kintsch, W. and Greeno, J.G.:*Understanding and solving word arithmetic problems*. Psychological Review, 1985, 92, 109-129.

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.

Polson, P.G. and Kieras, D.E.: *A quantitative model of the learning and performance of text editing knowledge*. Proceedings of the CHI 1985 Conference on Human Factors in Computing. San Francisco, April 1985.

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

Young, S. A theory and simulation of macrostructure. Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

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

The goal of the MENTOR project is to implement and begin evaluation of a computer-based methodology for reducing therapeutic misadventures. The project will use principles of artificial intelligence to create 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 will be automatically acquired through interfaces to a hospital information system. This data will be 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 will be fed back to the responsible physicians to assist future decision making.

Specific objectives of this proposal include:

1. Implement a prototype computer-based expert system to continuously monitor in-patient drug therapy. It will use 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 the proposed project will build 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.

Rutgers AIM Project: RUTGERS RESEARCH RESOURCE-COMPUTERS IN BIOMEDICINE

Principal Investigators: Casimir Kulikowski, Ph.D.
Sholom M. Weiss, Ph.D
Department of Computer Science
Rutgers University
New Brunswick, New Jersey 08903
(201) 932-2006 (KULIKOWSKI@RUTGERS)
(201) 932-2379 (WEISS@RUTGERS)

The Rutgers Research Resource provides the research support with artificial intelligence systems, and the computing support with its DEC2060 facility to a large number of biomedical scientists and researchers. Research activities are concentrated in two major areas: expert medical systems, models for planning and knowledge acquisition, and general AI systems development.

One of the most significant achievements in bringing the work of the Resource to bear on clinical research and practice lies in the transfer of technology from our large DEC20 machine to microprocessor compatible representations. The initial breakthrough came with the automatic translation of a serum protein electrophoresis interpretation model so that a version could be incorporated in an instrument - a scanning densitometer. It is now being used at several hundred clinical locations.

During the current period, we have been working on a new project with long term implications for the impact of AIM technology: the development of a hand-held microcomputer version of an expert consultation system for front-line health workers. In collaboration with Dr. Chandler Dawson (UCSF), Director of the World Health Organization's Collaborative Centre for the Prevention of Blindness and Trachoma, we have developed a prototype model for consultation on primary eye care. This has been oriented at problems of injury, infection, malnutrition and cataract in situations where an ophthamologist is unavailable. In most developing nations, the incidence of blindness is 10% to 40% higher than in the USA because of these kinds of problems. With the help of a grant from the USAID, we are developing the systems needed for management of eye disease by front-line health workers in developing nations, and outlying parts of the USA.

REFERENCES

Weiss, S.M. and Kulikowski, C.A. *A Practical Guide to Designing Expert Systems*, Rowman and Allanheld, 1984.

Kulikowski, C.A.*Expert Medical Consultation Systems*, Journal of Medical Systems, v.7, pp. 229-234, 1983.*

Weiss, S.M., Kulikowski, C.A., and Galen, R.S.,*Representing Expertise in a Computer Program: The Serum Protein Diagnostic Program*, Journal of Clinical Laboratory Automation, v.3, pp. 383-387, 1983.*

Kastner, J., Dawson, C., Weiss, S., Kern, K., Kulikowski, C., *An Expert Consultation System for Frontline Health Workers in Primary Eye Care*, Journal of Medical Systems, Vol. 8, No. 5 (1984).*

National AIM Project: SECS -- SIMULATION AND EVALUATION OF CHEMICAL SYNTHESIS

Principal Investigator: W. Todd Wipke, Ph.D.
Department of Chemistry
University of California at Santa Cruz
Santa Cruz, California 95064
(408) 429-2397 (WIPKE@SUMEX-AIM)

The SECS Project aims at developing practical computer programs to assist investigators in designing syntheses of complex organic molecules of biological interest. Key features of this research include the use of computer graphics to allow chemist and computer to work efficiently as a team, the development of knowledge bases of chemical reactions, and the formation of plans to reduce the search for solutions. SECS is being used by the pharmaceutical industry for designing syntheses of drugs.

A spin-off project, XENO, is aimed at predicting the plausible metabolites of foreign compounds for carcinogenicity studies. First, the metabolism is simulated; then the metabolites are evaluated for possible carcinogenicity.

SOFTWARE AVAILABLE ON SUMEX

No software is available on SUMEX after 31 March 1985 when this project left the SUMEX system. Contact Dr. Wipke directly.

REFERENCES

Wipke, W.T., Rogers, D.: *Rapid Subgraph Search Using Parallelism.* J. Chem. If. Comput. Sci., 24:4 255-262 (1984).

Wipke, W.T.: "An Integrated System for Drug Design" in *The Aster Guide to Computer Applications in the Pharmaceutical Industry* Aster Publishing Co., Springfield, Oregon, 1984, pp 149-166.

Wipke, W.T., and Rogers, D.: *Artificial Intelligence in Organic Synthesis. SST: Starting Material Selection Strategies. An Application of Superstructure Search.* J. Chem. Inf. Comput. Sci., 24:1 71-81, 1984.

Wipke, W.T.: *Computer Modeling in Research and Development*, Cosmetics and Toiletries, 99:Oct 73-82 (1984).

Wipke, W.T.: *Computer-Assisted Design of Organic Synthesis. ALCHEM: A Language for Representing Chemical Knowledge*, J. Chem. Info. Comput. Sci., 24, 0000 (1985).

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)

William B. Thompson, Ph.D.
Department of Computer Science
136 Lind Hall
University of Minnesota
Minneapolis, Minnesota 55455
(612) 373-0132 (THOMPSON@SUMEX-AIM)

The Minnesota SOLVER project focuses upon the development of strategies for discovering and representing the knowledge and skill of expert problem solvers. Although in the last 15 years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only 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.

SOFTWARE AVAILABLE ON SUMEX

A redesigned version of the Diagnoser simulation model, named Galen, has been implemented on SUMEX.

REFERENCES

Johnson, P.E., Moen, J.B., and Thompson, W.B.: *Garden Path Errors in Medical Diagnosis*. IN Bloc, L. and Coombs, M.J. (Eds.), COMPUTER EXPERT SYSTEMS, Springer-Verlag (in press).

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).

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.

Johnson, P.E.: *What kind of expert should a system be?* J. Medicine and Philosophy, 8:77-97, 1983.

Johnson, P.E., Duran, A., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P. and Swanson, D.: *Expertise and error in diagnostic* *reasoning*. Cognitive Science 5:235-283, 1981.

Thompson, W.B., Johnson, P.E. and Moen, J.B.: *Recognition-based diagnostic reasoning*. Proc. Eighth IJCAI, Karlsruhe, West Germany, August, 1983.

Stanford Pilot Project: THE COMPUTER-AIDED MEDICAL DECISION ANALYSIS (CAMDA) PROJECT

Co-Principal Investigator: Samuel Holtzman
Ronald A. Howard
Department of Engineering-Economic Systems
Stanford University
Stanford, California 94305

Contact: Samuel Holtzman(HOLTZMAN@SUMEX-AIM)
(415) 497-0486

The CAMDA project is a program of research in the area of medical decision making. The main focus of this effort is to combine decision analysis and artificial intelligence to develop systems that support medical decisions.

Nearly two decades of experience in the application of decision analysis to problems in industry and government have shown that the technique constitutes an extremely helpful tool for making difficult choices. The potential benefit of decision analysis is particularly great when choices must be made in the presence of uncertainty and when the stakes involved are high. This situation is common in medical decisions.

Partly as a result of the high cost of an individual decision analysis, and partly due to the inherent complexity of making choices which involve outcomes such as pain and death, medical decision analysis has remained essentially within the realm of the academic community. Therefore, the majority of patients and physicians have been deprived of the benefits of this powerful technique.

Expert system technology makes it possible to bring decision analysis to the medical community in general. By providing a sophisticated modeling methodology, expert systems allow the process of decision analysis (within a specific medical context) to be formalized with sufficient accuracy to make much of the analysis amenable to computer automation. The resulting CAMDA systems could provide an attractive alternative to unaided decision making, and to the usually unaffordable option of analyzing medical decisions individually. Furthermore, these systems can help decision makers think more clearly about the difficult issues they face by providing them with a means to experiment with the logical consequences of their assumptions and preferences.

A major focus of our research effort is the development of RACHEL, an intelligent decision system for infertile couples. The field of infertility was chosen for several reasons, including the prevalence of the condition, the complexity of the values that are usually attached to the possible outcomes in this field, the rapidly growing set of available tests and treatments, and the time-dependent nature of the human reproductive process.

As part of the development of RACHEL, a substantial portion of the current CAMDA effort is aimed at the development of a general computer-based aid for medical decision analysis, which could be used in other medical decision domains.

REFERENCES:

Holtzman, S.: *A Decision Aid for Patients with End-Stage Renal Disease,* Department of Engineering-Economic Systems, Stanford University, Stanford, California, 1983.

Holtzman, S.:*On the Use of Formal Methods for Decision Making,* Department of Engineering-Economic Systems, Stanford University, 1985.

(*) Holtzman, S.:*Intelligent Decision Systems,* Ph.D. Dissertation, Department of Engineering-Economic Systems, Stanford University, 1985.

Stanford Project: REFEREE Project

Principal Investigators:

Bruce G. Buchanan, Ph.D. (BUCHANAN@SUMEX-AIM)
Computer Science Department
701 Welch Road
Stanford University
Palo Alto, California 94304
(415) 497-0935

Byron W. Brown (BWROWN@SUMEX-AIM)
Department of Biostatistics
Stanford University Medical Center
Stanford, California 94305

Daniel E. Feldman (DFELDMAN@SUMEX-AIM)
Department of Medicine
Stanford University Medical Center
Stanford, California 94305

COLLABORATIVE PROJECT ABSTRACT

The goal of this project is two-fold: (a) use existing AI methods to implement an expert system that can critique medical journal articles on clinical trials, and (b) in the long term, develop new AI methods that extract new medical knowledge from the clinical trials literature. In order to accomplish (a) we are building the system in three stages.

1. System I will assist in the evaluation of the quality of a single clinical trial. The user will be imagined to be the editor of a journal reviewing a manuscript for publication, but the program will be tested on a variety of readers, including clinicians, medical scientists, medical and graduate students, and clerical help.

2. System II will assist in the evaluation of the effectiveness of the treatment or intervention examined in a single published clinical trial. The user will be imagined to be a clinician interested in judging the efficacy of the treatment being tested in the trial.

3. System III will assist in the evaluation of the effectiveness of a single treatment examined in a number of published clinical trials.

National AIM Project: Computer-Aided Diagnosis of Malignant Lymph Node Diseases (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.
Department of Medicine
Stanford University Medical Center - Room TC135
Stanford, California 94305
(415) 497-6979 (FAGAN@SUMEX-AIM)

We are building a computer program, called PATHFINDER, to assist in the diagnosis of lymph node pathology. The project is based at the University of Southern California in collaboration with the Stanford University Medical Computer Science Group. A pilot version of the program provides diagnostic advice on 80 common benign and malignant diseases of the lymph node 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.

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. Bennett, James S. and Robert S. Engelmore. "SACON: A Knowledge-Based Consultant for Structural Analysis." *IJCAI Proceedings 6* (1979), 47 - 49.

3. Bryson, Y.J., Dillon, M., Lovett, M., et al. "Treatment of first episodes of genital herpes simplex infection with oral acyclovir." *New England Journal of Medicine 308* (1983), 916-921.

4. Buchanan, B.G., and Shortliffe, E.H.. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Addison-Wesley Publishing Company, Reading, MA, 1984.

5. Coulter, C. L. "Research Instrument Sharing." *Science 201*, 4354 (1978).

6. Emerson, J.D., Graham, A.C. "Use of statistical analysis in the New England Journal of Medicine." *New England Journal of Medicine 309* (1983), 1332-1337.

7. Erman, L.D., Hayes-Roth, F., Lesser, V.R., and Reddy, D.R. "The Hearsay-II speech-understanding system: Integrating knowledge to resolve uncertainty." *Computing Surveys 12* (1980), 213-253.

8. Feinstein, A.R., Horwitz, R.I., Spitzer, W.O., Battista, R.N. "Coffee and Pancreatic Cancer." *JAMA 246* (1981), 957-961.

9. Genesereth, M.R. A Meta-Level Representation System. Tech. Rept. HPP83-28, Stanford University, Knowledge Systems Laboratory, 1983. The Report Consists of Four: HPP80-24, HPP82-27, HPP83-26, and HPP84-1

10. Haggerty, J. REFEREE and RULECRITIC: Two prototypes for assessing the quality of a medical paper. Master Th., Stanford University, 1984.

11. Hart, P.O., Duda, R.O., Einaudi, M.T. "PROSPECTOR -- A Computer-based Consultation System for Mineral Exploration." *Mathematical Geology 10*, 5 (1978).

12. Hayes-Roth, B. The blackboard architecture: A general framework for problem-solving? Tech. Rept. HPP-83-30, Stanford, Ca.: Stanford University, 1983.

13. Hayes-Roth, B. BB1: An architecture for blackboard systems that control, explain, and learn about their own behavior. Tech. Rept. HPP-84-16, Stanford, Ca.: Stanford University, 1984.

14. Hayes-Roth, B. "A blackboard architecture for control." *Artificial Intelligence Journal in press* (1985).

15. Hayes-Roth, F., Waterman, D. A. & Lenat, D. B. (Eds.). *Building Expert Systems.* Addison-Wesley, Reading, MA, 1983.

16. 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 17-14, 1984, 1984.

17. Kunz, J.C., Fallat, R.J., McClung, D.H., Osborn, J.J., Votteri, B.A., Nii, H.P., Aikins, J.S., Fagan, L.M., Feigenbaum, E.A. A physiological rule-based system for interpreting pulmonary function test results. Proceedings of Computers in Critical Care and Pulmonary Medicine, 1979, pp. 375-379.

18. Keith A. Lantz et. al. *V - System 4.1 Reference Manual.* December 1, 1983. Computer Systems Laboratory, Stanford University.

19. Lederberg, J. "Digital Communications and the Conduct of Science: The New Literacy." *Proceedings of the IEEE 66*, 11 (1978).

20. Levy, R.I. "Current status of cholesterol controversy." *Am J Med 74* (1983), 1-4.

21. Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A., and Lederberg, J.. *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project.* McGraw-Hill, New York, 1980.

22. MacMahon, B., Yen, S., Trichopoulos, D., et al. "Coffee and cancer of the pancreas." *New England Journal of Medicine 304* (1981), 630-633.

23. Masinter, Larry M. Interlisp VAX: A Report. Tech. Rept. HPP-81-14 or STAN-CS-81-879, Heuristic Programming Project, Stanford University, 1981.

24. McDermott, J. "R1: A rule-based configurer of computer systems." *Artificial Intelligence 9*, 1 (1982), 39-88.

25. Nii, H.P. and Aiello, N. AGE: A Knowledge-based Program for Building Knowledge-based Programs. Proceedings of IJCAI 6, 1979, pp. 645-655.

26. Nilsson, N.. *Principles of Artificial Intelligence.* Tioga, Palo Alto, CA, 1980.

27. Novak, Gorden. *GLISP User's Manual.* Heuristic Programming Project, Stanford University, Stanford Ca. 94305, 1982.

28. Price, D.S. The development and structure of the biomedical literature. In *Coping with the Biomedical Literature*, Praeger, New York, 1981.

29. Rich, E.. *Artificial Intelligence.* McGraw-Hill, New York, 1983.

30. Rout, L. "Fast medical advances give doctors troubles in staying up to date." *Wall Street Journal 104* (1981), 1, 15.

31. Sackett, D.L. "How to read clinical journals: I. Why to read them and how to start reading them critically." *Can Med Assoc J 124* (1981), 555-558.

32. Shortliffe, Edward H.. *Computer-Based Medical Consultation: MYCIN.* American Elsevier, New York, 1976.

33. Smith, Reid G., and Friedland, Peter. Unit Package User's Guide. Tech. Rept. KSL-80-28, Knowledge Systems Laboratory, Stanford University, December, 1980.

34. Van Melle, W., Scott, A.C., Bennett, J.S., and Peairs, M.A.S. The EMYCIN Manual. Tech. Rept. KSL-80-11, Knowledge Systems Laboratory, Stanford University, 1980. Superseded by KSL-81-16

35. Winston, P.. *Artificial Intelligence, 2nd ed.* Addison-Wesley, Reading, MA, 1984.