Investigators and are implemented by direct charges against designated user research grants and contracts. This avoids the large accounting overhead of a cost center to collect relatively small bills each month.

III.E. Dissemination of Resource Information

We have continued our past substantial efforts to disseminate the AI technology developed at SUMEX-AIM. This has taken the form of many publications -- over forty-five combined books and papers are published per year by the KSL; wide distribution of our software, including systems software and AI application and tool software, both to other research laboratories and for commercial development; production of films and video tapes depicting aspects of our work; special seminars to introduce users to the systems we develop; and significant project efforts at studying the dissemination of individual applications systems such as the ONCOCIN resource-related research project. We continue to provide active support for the AIM workshops. The most recent one was held in the spring of 1988 at Stanford University, under the auspices of the American Association for Artificial Intelligence (AAAI). Planning is underway for an AIM workshop in the spring of 1990.

III.E.1. Software Distribution

We have widely distributed both our system software and our AI tool software. Since much of our general system-level software is distributed via the ARPANET we do not have complete records of the extent of the distribution. Software such as TOPS-20 monitor enhancements, the Ethernet gateway and TIP programs, the SEAGATE AppleNet to Ethernet gateway, the PUP Leaf server, the SUMACC development system for Macintosh workstations, and our Lisp workstation programs are frequently distributed in this manner to the ARPANET community and beyond.

Our primary distribution effort is directed towards the AI tools we have developed. In recent years, the volume of inquiries for this type of software and requests for tapes has been a substantial burden on the staff, especially in the face of recent budget cuts. As a result, we have turned over most of this type of software distribution to Stanford's Office of Technology Licensing (OTL). This organization handles software distribution and technology licensing matters for much of the Stanford community. Since there are several OTL staff members assigned to the distribution of Stanford software, requests for information and tapes are handled quickly and efficiently. Also, OTL's staff has the expertise needed to handle the legal questions that frequently arise in the distribution of software, and has an established computerized record-keeping scheme. The SUMEX staff continues to be available as needed to assist OTL with special administrative and technical matters.

Distribution continues for the Parallel Computing Architectures Project multiprocessor simulation system, CARE/SIMPLE, the EMYCIN package, and the BB1 package.

III.E.2. AIM Community Systems Support

We continue to make a special effort to assist other members of the SUMEX-AIM community in integrating the technologies needed for biomedical AI research. This is often achieved through direct contact by SUMEX staff members with researchers at these institutions, (e.g., with Professor Mark Frisse's and Michael Kahn's groups at Washington University, Dr. James Brinkley's group at the University of Washington, and Professor Widman's group at the University of Texas), at meetings and workshops, or via electronic mailing lists. For example, the Info-MAC, Info-Explorer, and Info-1100 mailing lists have hundreds of members and cover a broad range of equipment issues, software issues, and topics in artificial intelligence.

III.E.3. Video Tapes and Films

Various groups in the KSL have continued to prepare video tapes that provide an overview of the research and 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. In addition to the earlier tapes covering Knowledge Engineering in the KSL, ONCOCIN Overview, and ONCOCIN Demonstration, we have recent tapes on the PROTEAN project, the BB1 project, and a one-day symposium on KSL research activities.

III.E.4. Special Seminars

SIMPLE/CARE is a powerful simulation system which permits empirical studies of expert system performance on a wide class of multicomputer architectures, including quantitative measurements of system behavior. Our simulation system is now in use by several research groups at Stanford, and it has been ported to several external sites, including NASA Ames Research Center. A videotaped tutorial was held in June, 1988, attended by representatives from industry and government, which described the CARE/SIMPLE system, as well as the LAMINA programming interface . The attendees received instruction in use of the system for making measurements of the performance of various simulated multiprocessor applications.

Due to rapidly growing interest in the SIMPLE/CARE system, a major effort is now underway to port it to wider class of hardware platforms. The system is currently being reimplemented in Common Lisp and the X window system, with the Sun workstation as the initial target.

III.F. Suggestions and Comments

III.F.1. 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.

III.F.2. 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 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. This is particularly important in the light of significant on-going changes to the national networking environment.

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 resource at Stanford. Many groups from the National AIM community now use the SUMEX-AIM resource solely for communication (i.e., electronic mail to and from colleagues or access to bulletin boards and other information resources at SUMEX). Because of the difficulty of recording Internet connections and system-level mail forwarding and related communications services, we can no longer accurately keep a list of these users. However, from the usage data shown in Section III.A.2.8, the volume of these services continues to rise as the AIM community moves increasingly to distributed resources.

The detailed 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
     1. Accomplishments this past year
     2. 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
     1. Near-term
     2. 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 Shortliffe as Principal Investigator.

IV.A.1. Guardian Project

Project Leader: Barbara Hayes-Roth, Ph.D.
Department of Computer Science
Stanford University

Collaborator: Adam Seiver, M.D.
Department of Surgery
Veterans Administration Hospital
Palo Alto, CA.

I. Summary of Research Program

A. Project Rationale

Critical care depends upon sophisticated life-support technology. Devices such as the respirator, dialysis machine, and intra-aortic balloon pump maintain life until the patient's own organs heal and resume normal function. However, effective management of device-supported patients is complex, involving interpretation of a large number of physiological variables, comparative evaluation of multiple therapeutic options, and control of many device parameters. Even skilled clinicians can make errors that produce life-threatening situations or otherwise harm the patient. These problems are compounded when the number of patients requiring life-support technology exceeds the availability of skilled clinicians.

Short-term research on computer-based assistance for critical care aims to alleviate some of these problems. Research on "smart alarms" aims to improve capabilities for signaling abnormal patient data, while reducing the false alarm rates associated with current alarm systems. Research on automatic tracking and combining of patient data values aims to help clinicians identify a wider range of problems. Research on display technology aims to help clinicians select, examine, and interpret important patient data. However, these short-term approaches do not address the fundamental problem of effectively managing an increasingly complex and sophisticated life-support technology.

From a longer-term perspective, what is needed is an "intelligent" computer system that integrates knowledge of underlying causal mechanisms and knowledge of the broader patient context--knowledge that currently is distributed among different experts on the critical care team. Such a system would acquire patient data automatically, synthesize data into a dynamic model of the patient's physiological functioning, and dynamically plan effective programs of device settings and other therapeutic actions. Acting in the role of an intelligent critical-care consultant, it would explain its observations, reasoning, conclusions, and recommendations to clinical care staff. Working toward this longer-term objective, we are developing "Guardian," a prototype system for intelligent patient monitoring in the surgical intensive care unit (SICU).

Development of Guardian poses many challenging research problems, including the following. How should we represent structure/function knowledge of biological systems and life-support devices? How should we represent the time-varying course of component biological processes and of the patient's medical condition? What kind of architecture will allow Guardian to integrate signal processing, knowledge-based reasoning, and user interaction? How can Guardian monitor a growing number of patient data parameters, many of them sampled several times per second, with limited computational resources? How should Guardian perform each of its component reasoning tasks? How can Guardian control its reasoning behavior--choice of reasoning tasks and strategies for performing particular tasks--so as to meet real-time constraints on the utility of its conclusions? These questions provide foci for the research of individual students on the project.

B. Medical Relevance and Collaboration

The proposed research aims to bring about a fundamental advance in the medical device technology base. Life-support devices are an accepted foundational element of critical care and they continue to advance in power and sophistication. However, like much advanced equipment, increasingly sophisticated life-support devices contribute to an increasingly complex information processing task for the people who use them. In the best of circumstances, cognitive overload threatens to undermine the utility of these devices and the quality of critical care. Short-term research directed at smart alarms, low-level data interpretation, and effective data displays will produce useful products, but it will not solve this long-term problem.

The proposed research attacks the long-term problem of effective critical care management with the methods of artificial intelligence. In essence, we aim to capture the knowledge and skills of critical care experts in a compute program that could assist skilled clinicians or stand in for unavailable staff. Thus, our goal is to create a device-management technology that is every bit as powerful as the device technology itself. We believe that this is the only way to fully exploit evolving life-support technology and insure high quality patient care.

In particular, an intelligent critical care consultant potentially could provide three kinds of assistance:

First, it could enhance the management and care of severely injured persons. It could guarantee the availability of expertise held by different medical experts (e.g., surgeon, nurse, physician specialists) at the time it is needed. It could provide continuous and vigilant patient monitoring. It could help attending clinicians to notice easily overlooked events, to analyze problems in sufficient detail, to consider alternative interpretations and treatments, and to avoid making errors.

Second, it could reduce workload and facilitate care tasks for providers. It could perform routine patient monitoring and provide summary accounts for

analysis and decision-making by clinicians. It could substitute for physician specialists whose expertise is required elsewhere. It could assume responsibility for managing stable patients, freeing available clinicians to focus on more seriously injured patients.

Third, it could provide continuing training and consultations to students or less well trained critical care staff.

Because the Guardian project depends upon state-of-the-art artificial intelligence methods, as well as extensive and sophisticated medical knowledge, it is being conducted as a collaboration between Dr. Barbara Hayes-Roth and Dr. Adam Seiver. Dr. Hayes-Roth designed the BB1 dynamic control architecture, which is the foundation for Guardian and several other applications involving real-time monitoring and control in other domains. Dr. Seiver is Associate Director of the new 14 bed surgical intensive care unit at the Palo Alto Veterans Administration Medical Center. At this time, the project also includes two post-doctoral fellows, Dr. Rattikorn Hewett and Dr. Luc Boureau, as well as three Ph.D. students, Mr. Richard Washington, Mr. Adnan Darwiche, and Mr. David Ash. In addition, we cooperate informally with Dr. Lawrence Fagan and his students, who are involved in the VentPlan project, which takes a control theoretic approach to complementary aspects of SICU monitoring, primarily the moment-by-moment optimization of device settings.

C. Research Progress

Following a period of task analysis, knowledge acquisition, and conceptual design, we developed the first version of Guardian in the winter of 1988. Since then, we have iterated a series of research cycles involving: (a) conceptual development of component approaches to knowledge representation and particular reasoning tasks; (b) implementation and integration of component approaches in a new version of Guardian; (c) demonstration of the new version on an important SICU scenario; and (d) solicitation of feedback and criticism from knowledgeable colleagues. Each such cycle takes about two academic quarters.

The current version of Guardian integrates these reasoning components: FOCUS preprocesses — abstracts, classifies, filters, and rates for urgency and importance — sensed data before relaying them to the reasoning system. FOCUS currently performs these functions for about 25 continuously monitored patient-data parameters. ASSOC-REACT uses probabilistic associations to diagnose commonly occurring problems and to suggest standard treatments. Recent innovations involving hierarchical organization of knowledge permit ASSOC-REACT to circumvent the well-known computational complexity of its underlying belief network mechanism and, in fact, to modulate its processing time according to the time available. ICE uses explicit structure/function models of biological systems — currently the respiratory, circulatory, and tissue metabolism systems — to suggest plausible underlying causes for unfamiliar or otherwise inexplicable

conditions. Recent innovations permit ICE to predict the consequences of observed or hypothetical conditions as well. TPLAN suggests courses of therapy actions over time to deal with evolving medical conditions. Each of these reasoning components is implemented in a domain-independent fashion and can be applied to any biological or other system for which the relevant knowledge is available.

D. Demonstrations, Presentations, and Publications

We have given demonstrations of the Guardian system to many colleagues in the medical AI and larger AI communities, for example to: Dr. Lawrence Fagan and his students from Stanford's Medical Computing Systems Group; Dr. Seppo Kalli, Director of Medical Signal Processing at Technical Research Centre of Finland; Dr. William Pardee and his associates from the Rockwell Science Center; Dr. Perry Thorndyke and his associates from FMC Corporation; Dr. Joseph Naser of the Electric Power Research Institute.

We have made (or plan to make) invited presentations of this research to: IEE Workshop on Expert Control Systems, Brighton England, June, 1990 International Joint Conference on Artificial Intelligence, Detroit, Aug, 1989 AI Systems in Government Conference, Washington D.C., March, 1989 AAAI Symposium on Knowledge System Development Tools, Stanford, March, 1989 Stanford SIGLunch, February, 1989 Workshop on Formal Aspects of Semantic Networks, Catalina, February, 1989 Carnegie Symposium on Architectures for Intelligence, Pittsburgh, May, 1988 Advanced Decision Systems, Palo Alto, May, 1988 Boeing Computer Services, Bellevue, WA., March, 1988 DARPA Knowledge-Based Planning Workshop, Austin, December, 1987

The project has produced the following publications:

- Hayes-Roth, B., Washington, R., Hewett, R., Hewett, M., and Seiver, A. Intelligent monitoring and control. Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI-89, Detroit, Mi., 1989.
- Washington, R., and Hayes-Roth, B. Input data management in real time AI systems. Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI-89, Detroit, Mi., 1989.
- Hayes-Roth, B. Making intelligent systems adaptive. In K. VanLehn (Ed.) Architectures for Intelligence. Lawrence Erlbaum, 1989.
- Hayes-Roth, B., Hewett, M., Washington, R., Hewett, R., and Seiver, A. Distributing intelligence within a single individual. In L. Gasser and M.N. Huhns (Eds.) Distributed Artificial Intelligence Volume 2. Morgan Kaufmann, 1989.
- Hewett, R., and Hayes-Roth, R. Representing and reasoning about physical systems using generic models. In J. Sowa (Ed.) Formal Aspects of Semantic Networks. Morgan Kaufmann, 1989.

- Hayes-Roth, B. Dynamic control planning in adaptive intelligent systems. Proceedings of the DARPA Knowledge-Based Planning Workshop, 1989.

IV.A.2. MOLGEN Project

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

Prof. E. Feigenbaum
Department of Computer Science
Stanford University

Dr. P. Friedland
NASA-Ames Research Center
Moffett Field, CA

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 research has taken the specific form of systems that provide assistance to the experimental biologist in various tasks, including the design of complex experiment plans, the analysis of nucleic acid sequences, and the formulation of hypotheses in the subdomain of regulatory genetics.

B. Medical Relevance and Collaboration

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

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

needs of over two thousand academic molecular biologists in the United States.

Our recent work in using qualitative reasoning techniques to encode theories of molecular biology is likely to be relevant to the human genome project. We have constructed models of the structure and function of the tryptophan operon gene-regulation system, including its enzymatic pathways, gene-regulation mechanisms, and the general processes involved in gene expression such as transcription and translation. The methods we developed to model this bacterial system will be applicable to modeling both the regulation of the thousands of genes that will be sequenced in the human genome project, and the activities of the protein products of these genes. Our work in hypothesis formation will aid molecular biologists in formulating to explain such processes as gene regulation, by referring to such large knowledge bases of biological knowledge.

C. Highlights of Research Progress

C.1 Accomplishments

The MOLGEN project has successfully concluded with the publication of Peter Karp's doctoral dissertation. Here we summarize the contributions of that dissertation.

Karp's dissertation investigates scientific reasoning from a computational perspective. The investigation focuses upon a program of research in molecular biology that culminated in the discovery of a new mechanism of bacterial gene regulation, namely Dr. Charles Yanofsky's discovery of attenuation. In the first phase of this research, the MOLGEN group performed a historical study of the biological research that reconstructed the different theories that the biologists possessed at different points in time, and analyzed the differences between successive theories. In the second phase, Karp developed a qualitative chemistry for representing theories of molecular biology. In the third phase of the research, he constructed a computer program that solves hypothesis-formation problems encountered by the biologists.

C.1.1 Qualitative Modeling and Simulation

In order to solve the hypothesis-formation task, we must have a framework for representing theories in molecular biology that allows those theories to be used to predict the outcomes of laboratory experiments. The representation must also allow the hypothesis-formation program to reason about a theory and modify it in order to improve the predictive power of the theory. Chapter 3 of the dissertation presents three related methods for representing scientific theories. The third representation method was selected for use in conjunction with the hypothesis-formation task. This method breaks theories in molecular biology into several parts: A *class knowledge base* defines a taxonomic hierarchy of the classes of biological objects that exist in the trp operon. A *process knowledge base* describes the chemical reactions that can

occur between biological objects. An experiment is described in a third knowledge base by creating the particular objects (instantiated from the known classes of objects) that are present in the experiment.

A qualitative reasoning program called GENSIM determines what reactions occur between the objects in an experiment; these reactions create new objects, which can cause additional reactions. Chapter 3 shows that a process-oriented framework for qualitative simulation is more flexible than a framework based on a fixed network of the state variables of a physical system. The GENSIM program defines a qualitative chemistry --- a framework for reasoning about chemical reactions --- and identifies constraints that a chemical modeling system must satisfy to simulate chemical reactions correctly. In addition, the chapter presents new qualitative representations for capturing the partial knowledge that biologists (and other scientists) have of the mathematical relationships that characterize the systems they study.

C.1.2 A Historical Study of the Discovery of Attenuation

Chapter 4 contains a detailed historical study of the process by which Dr. Yanofsky and his colleagues discovered attenuation. The study is based on information obtained from the scientific publications the biologists produced, and from interviews with the biologists. This biological research program consumed over 50 man-years of effort, and so is one of the most complex ever studied in Artificial Intelligence.

In the first phase of the analysis, the MOLGEN group produced a conceptual reconstruction of what knowledge the biologists possessed about the trp operon at different points in time. In the next phase, Karp searched for patterns in the differences between successive states of the biologists' knowledge. These differences were due to changes the biologists made to their theories of the trp operon. Patterns in the differences indicate reasoning methods that were used to derive new theories from old. This analysis suggests that biologists use *theory modification operators* to modify a theory such that its predictions are altered. These patterns also support the conjecture that scientists use four different *modes of scientific exploration* to determine what types of experiments to perform next from a given state of research. A mode of exploration selects experiments based on the number of theories entertained at a given moment, and their relative credibilities.

C.1.3 Hypothesis Formation

Chapter 5 describes methods for solving the hypothesis-formation problem. The problem is to generate hypotheses that rectify a discrepancy between the observed outcome of an experiment, and the outcome predicted by GENSIM. A hypothesis modifies either GENSIM's theory or the initial conditions of the experiment (which are often not known with certainty), such that the predicted outcome of the experiment matches its observed outcome. The thesis treats the problem of hypothesis formation as a *design problem*. The goal of the designer is to eliminate the difference between the observed and

predicted outcomes of the experiment -- the *prediction error*. A hypothesis is synthesized by *design operators* that reason backward from the prediction error and determine what modifications to the theory or initial experimental conditions will eliminate the error. HYPGENE uses a sophisticated planning system to perform backward reasoning; the planner can achieve goals represented as arbitrary predicate-calculus formulae, including quantification.

This design problem is a search problem because often more than one operator is relevant to eliminating a prediction error, and a single operator can sometimes be applied in several ways. The synthetic, goal-directed search used here should prove more efficient than past approaches to hypothesis generation, which often used heuristic search to guide a purely syntactic generator of hypotheses. HYPGENE uses heuristic search to guide a generator that is already focused on errors in the prediction. Its search can be guided further by the results of a second, similar experiment which we term a *reference experiment*. The difference between the initial conditions of the two experiments is likely to have caused the prediction error, and can be used evaluate hypotheses HYPGENE has generated. In addition, the thesis describes a reasoning mode which can generate hypotheses by reasoning forward from the difference in initial conditions. This forward reasoning mode is more efficient than the backward mode for some problems.

Chapter 7 is an empirical investigation of the methods in Chapters 3 and 5, in which GENSIM and HYPGENE were run on several test cases. GENSIM predicted the outcomes of several biological experiments. HYPGENE formulated hypotheses to account for several incorrect GENSIM predictions. Most of these hypothesis-formation problems were taken from the historical study of the biologists research; HYPGENE produced many of the same hypotheses as the biologists did. The chapter summarizes the strengths and weaknesses of the methods, as revealed by these tests.

HYPGENE provides a *flexible* framework for hypothesis formation because its framework is syntactically complete -- its operators can modify the initial conditions of the experiment, the process knowledge base, or the class knowledge base. HYPGENE's flexibility is also enhanced because its planner can manipulate complex predicate-calculus expressions, allowing the program to reason about complex domain processes. Because HYPGENE's planner and operators do not contain domain concepts, the framework is largely domain independent. The framework is efficient because HYPGENE's planner works backward from prediction errors using operators that associate syntactic classes of prediction errors with specific types of theory modifications. The framework allows us to integrate domain-specific knowledge (such as general knowledge of chemistry) into the hypothesis generator to prune partial solutions during the generation process. Efficiency is further increased by the use of reference experiments, which provide information for both filtering and generating hypotheses.

D. Publications

1. Bach, R., Friedland, P., and Iwasaki, Y.: *Intelligent computational assistance for experiment design*. Nucleic Acids Res. 12(1):11-29, January, 1984.
2. Friedland, P., and Kedes, L.: *Discovering the secrets of DNA*. Communications of the ACM, 28(11):1164-1186, November, 1985, and IEEE/Computer, 18(11):49:69, November, 1985.
3. Friedland, P. and Iwasaki Y.: *The concept and implementation of skeletal plans*. Journal of Automated Reasoning, 1(2): 161-208, 1985.
4. Friedland, P., Armstrong, P., and Kehler, T.: *The role of computers in biotechnology*. BIO\TECHNOLOGY 565-575, September, 1983.
5. Karp, P., and D. Wilkins: *An Analysis of the Deep / Shallow Distinction for Expert Systems*. To be published in International Journal of Expert Systems, 1988.
6. Karp, P., and P. Friedland: *Coordinating the Use of Qualitative and Quantitative Knowledge in Declarative Device Modeling*, in Artificial Intelligence, Simulation, and Modeling edited by L. Widman and K. Loparo, 1989.
7. Karp, P.: *A Process-Oriented Model of Bacterial Gene Regulation*, Stanford University Knowledge Systems Laboratory Technical Report KSL-88-18.
8. Round, A.: *QSOPS: A Workbench Environment for the Qualitative Simulation of Physical Processes*. Stanford University Knowledge Systems Laboratory Report KSL-87-37, 1987.
9. Karp, P.: *Hypothesis Formation by Design*, in Computational Models of Scientific Theory Formation, edited by J. Shrager and P. Langley, 1989, in press.
10. Karp, P.: *Hypothesis Formation and Qualitative Reasoning in Molecular Biology*, PhD Dissertation, Stanford University Computer Science Department, 1989.
11. Meyers, S. and Friedland, P.: *Knowledge-based simulation of regulatory genetics in bacteriophage Lambda*. Nucleic Acids Res. 12(1):1-9, January, 1984.

E. Funding Support

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

provided by the Defense Advanced Research Projects Agency, under contract N00039-86C-0033.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

SUMEX-AIM continued to serve as the nucleus of our computing resources. The facility not only provided excellent support for our programming efforts, but served as a major communication link among members of the project. Systems available on SUMEX-AIM such as EMACS, MM, Scribe and BULLETIN BOARD have made possible the project's documentation and communication efforts.

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

III. RESEARCH PLANS

This project has concluded successfully with the completion of Dr. Karp's thesis; no future research is planned at this time.

IV.A.3. ONCOCIN Project

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

Project Director: Lawrence M. Fagan, M.D., Ph.D
Department of Medicine
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

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

B. Medical Relevance and Collaboration

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

C. Highlights of Research Progress

C.1.A Background and Overview of Accomplishments

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

The first version was a character-based system run on the large DEC mainframes. Because of the limitations in the display capabilities and the requirement to be tethered to a time-shared mainframe computer, we began a re-implementation of the system in 1984. In 1985, we discontinued the mainframe version of ONCOCIN. The performance of the mainframe version

of ONCOCIN was documented in two evaluation papers that appeared in clinical journals (see Hickam and Kent's papers).

In 1986, we placed the workstation version of ONCOCIN into the Oncology Day Care clinic. This version is a completely different program from the version of ONCOCIN that ran on the DECsystem 20--using protocols entered through the OPAL program, with a new graphical data entry interface, and a revised knowledge representation and reasoning component.

The process of entering a large number of treatment protocols in a short period of time led to other research topics including: design of an automated system for producing meaningful test cases for each knowledge base, modification of the design and access methods for the time-oriented database, and the development of methods for graphically viewing multiple protocols that are combined into one large knowledge base.

In 1987, we began to explore the use of continuous speech recognition as an alternate entry method for communicating with ONCOCIN. This project requires the connection of speech recognition equipment produced by Speech Systems, Inc. of Tarzana to the ONCOCIN interface module. Christopher Lane has developed a prototype network connection and command interpreter between the speech module and the Xerox 1186 computer that runs ONCOCIN. Clifford Wulfman has designed a series of modifications to the ONCOCIN user interface to allow for verbal commands. This work is described in more detail in the core ONCOCIN section.

The majority of our effort during the last two years has been to understand the limitations of the clinic version of ONCOCIN, and to concentrate on the generalization of these techniques to other application areas besides oncology. The majority of this research is thus described as part of the core research discussion on ONCOCIN. Highlights of this year include: (1) development of a general knowledge acquisition tool (PROTÉGÉ) designed to handle skeletal planning applications for clinical trials in any area of medicine, (2) demonstration that the therapy planning and knowledge acquisition tools for ONCOCIN can be closely integrated, and (3) development of a speech input system for ONCOCIN.

As a demonstration of the capabilities of the project to date, we undertook an experiment to see how difficult and time-consuming it is to bring up a new treatment protocol. A summary of a recent colon protocol was down-loaded from the PDQ protocol database. Approximately 60% of the knowledge of the protocol summary fit easily into the OPAL high level description. Additional rules were entered using lower level editors. A limited consultation was run after about 4 hours of work. Although this is only one data point, we believe that it validates the generality of the knowledge acquisition and therapy planning approach that we have pursued for nearly a decade. Work continues on extending the knowledge acquisition and therapy planning tools to allow for a higher percentage of concepts that can be entered with the smallest possible amount of low level Lisp changes.

Although we have completed the transfer of ONCOCIN into a stable and useful system on the Xerox Lisp workstations, it is now clear that this type of machine will not provide the type of dissemination hardware we would like to see. There are no planned additions to increase the speed, decrease the cost, or increase the integration capabilities of these workstations. Although there may be other solutions that will allow us to port ONCOCIN directly to alternative hardware platforms, we need to move away from Xerox workstations and InterLisp language upon which most of our software is based. We are particularly interested in exploring the Mac II hardware. During the last six months, we have begun an experiment to port ONCOCIN to a TI Explorer board inside of a Mac II. We have completed the translation of the Ozone object-oriented system, the temporal network and most of the reasoner. We will next approach the design of the user interface, which must be rewritten anew, since the current interface depends heavily on the graphical capabilities of the Xerox workstations.

C.1.B Review of Research Issues in ONCOCIN and OPAL

Our work to refine the clinic versions of ONCOCIN and OPAL reached a mature stage during this last research year. As our attention has moved to the generalization of these tasks (E-ONCOCIN and PROTÉGÉ) it seems appropriate to describe the range of research issues that we have examined during the development of the ONCOCIN system.

Research Issues in the Development of the ONCOCIN Reasoner and Interviewer

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

  Our experience with adding a large number of protocols has led to the evaluation of the design of the internal structure of the knowledge base (e.g., the way we describe the relationships between chemotherapies, drugs, and treatment visits). We will continue to improve the method for traversing the plan structure in the knowledge base, and consider alternative arrangements for representing the structure of chemotherapy plans. Currently, the knowledge base of treatment guidelines and the

patient database are separated. We propose to tie these two structures closer together. Additional work is anticipated on turning ONCOCIN into a critiquing system, where the physician enters their therapy and ONCOCIN provides suggestions about possible alternatives to the entered therapy. Although we have concentrated our review of the ONCOCIN design primarily on the data provided by additional protocols, we know that non-cancer therapy problems may also raise similar issues. The E-ONCOCIN effort is designed to produce a domain-independent therapy planning system that includes the lessons learned from our oncology research. Samson Tu is primarily responsible for continued improvement of the reasoning component of ONCOCIN.

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

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

- *System support for the reorganization.* The LISP language, which we used to build the first version of ONCOCIN, does not explicitly support basic knowledge manipulation techniques (such as message passing, inheritance techniques, or other object-oriented programming structures). These facilities are available in some commercial products, but none of the existing commercial implementations provide the reliability, speed,

size, or special memory-manipulation techniques that are needed for our project. We have therefore developed a "minimal" object-oriented system to meet our specifications. The object system is currently in use by each component of the new version of ONCOCIN and in the software used to connect these components. In addition, all ONCOCIN student projects are now based on this programming environment. Christopher Lane created and is responsible for modifications to the object-oriented system.

Interactive Entry of Chemotherapy Protocols by Oncologists (OPAL)

The OPAL system permits physicians who are not computer programmers to enter protocol information on a structured set of forms presented on a graphics display. Most expert systems require tedious entry of the system's knowledge. In many other medical expert systems, each segment of knowledge is transferred from the physician to the programmer, who then enters the knowledge into the expert system. We have taken advantage of the generally well-structured nature of cancer treatment plans to design a knowledge entry program that can be used directly by clinicians. The structure of cancer treatment plans includes:

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

Using the graphics-oriented workstations, this information is presented to the user as computer-generated forms which appear on the screen. After the user fills in the blanks on the forms, the program generates the rules used to drive the reasoning process. As the user describes more detailed aspects of the protocol, new forms are added to the computer display; these allow the user to specify the special cases that make the protocols so complicated. Although the user is unaware of the creation of the knowledge base from the interaction with OPAL, a complex set of translations are taking place. The user's entries are mapped into an intermediate data structure (IDS) that is common for all protocols. From the IDS, a translation program generates rules for creating and modifying treatment, and integrates them with the existing ONCOCIN knowledge base. Considerable effort has been expended on producing a standard relational database as the appropriate data structure to underlie the OPAL IDS. The PROTÉGÉ system described in the core ONCOCIN section was built upon this relational database.

Although the "forms" were specifically designed for cancer treatment plans, the techniques used to organize data can be extended to other clinical trials, and eventually to other structured decision tasks. The key factor is to exploit the regularities in the structure of the task (e.g., this interface has an extensive notion of how chemotherapy regimens are constructed) rather than to try to build a knowledge-entry system that can accept *any* possible problem

specification. The OPAL program is based upon a domain-independent forms creation package designed and implemented by David Combs. This program will provide the basis for our extension of OPAL to other application areas.

We have now entered thirty-five protocols covering many different organ systems and styles of protocol design. Based on this experience, we continue to explore ways to modify OPAL to increase the percentage of the protocol that can be entered directly by our clinical collaborators. One direction in which we have extended the OPAL program is in providing a graphical interface of nodes and arcs to specify the procedural knowledge about the order of treatments and important decision points within the treatments. This work is described in several papers by Musen.

C.2 Research in Progress

The major thrusts in speech input and generalized knowledge acquisition are described in the core research description of ONCOCIN. We will describe here our research in complex therapy planning and it's spin-offs in temporal representations and summarization of patient records.

C.2.1 Strategic Therapy Planning (ONYX)

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

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

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

A major spin-off from our ONYX work is a program that can summarize temporal trends in patient visits during chemotherapy and produce a

summary of the patient's course using both data from the flowsheet and an *underlying model of bone marrow physiology*. This work has led to major improvements in the temporal representation and in the integrate of mathematical and symbolic models. This work is part of Michael Kahn's Ph.D. thesis.

Summarization is defined as the task of combining multiple observations or features into a more general statement and abstraction as the task of selecting a subset of available features considered most relevant to answering a particular question. Both tasks require a model of the underlying system that encodes extensive knowledge about the entities and relationships that cause the system behavior and result in the observations. In the setting of a dynamic system, the model must be capable of representing temporal relationships between entities.

This work proposes that the combination of mathematical and symbolic techniques can be used to construct useful summaries of complex time-ordered data. In particular, mathematical models are used to capture the knowledge about the physiological processes that are responsible for the patient's clinical findings. Model parameters represent physiological concepts that are clinically relevant for medical problem solving. Prior to any patient-specific observations, the model parameters are set to population-based estimates. Standard curve-fitting techniques using a Bayesian updating scheme adjust model parameters to new observations. As more patient-specific observations are obtained, the set of estimated model parameters move further away from the population estimates. Symbolic models are used to augment the mathematical model parameter and state estimates. As the patient's clinical course evolves, the symbolic model captures the concurrent contexts that affect the interpretation of the physiological model results. For example, a heart rate of 120 is considered abnormally high in the context of a resting person but may be inappropriately low in the context of a treadmill stress test. A key feature of the combined mathematical and symbolic approach is that the physiological model changes over time as additional data are obtained and the symbolic model modifies the interpretation of these model changes in light of the clinical contexts present when the data was observed.

The methodology for combining mathematical and symbolic models emphasizes four main elements in summarizing complex time-ordered data:

1) A mechanistically-motivated model (in medicine, a physiological model) forms the basis for converting raw observations into more meaningful concepts. However, the interpretation of these concepts requires additional knowledge such as the contextual information contained in a symbolic model.
2) The initial model is based on general knowledge since no specific observations are available to alter the initial impression. New observations will change the initial model by incorporating the new

information. The collection of altered models captures state changes that have evolved over time.

3) Differences in key model features or states form the basis for selective abstraction and effective summarization. A method for determining which features are pertinent to a user question or sufficiently "interesting" to warrant inclusion into a summarization requires additional domain-specific reasoning.

4) The construction of a concise and useful summarization requires the use of additional contextual and domain-specific information so that the generated summary text conforms to the user's expectations and requirements.

These principles form the basis for a computer program designed to summarize the clinical course of individual patients receiving experimental cancer chemotherapy. In this setting, patients are often receiving more than one treatment that have overlapping schedules and durations of action. Thus our temporal model requires the representation and the reasoning with multiple, simultaneous contexts to ensure the proper interpretation of a given observation or model estimate. ONCOCIN uses a specialized structure called the temporal network to represent treatment contexts used in temporal queries into a time-oriented patient record. We have extended the temporal network concept to create a symbolic model of the patient's clinical course over time. This structure permits the representation of multiple, concurrent contexts over time and therefore can capture the complex temporal nature of our patient's clinical course. For the proper interpretation of the mathematical model output, the temporal network provides the set of contexts that existed when the observation and model estimates were obtained. In addition, the interpretation task requires complex context-sensitive reasoning. For example, the interpretation of a model parameter may be different if two contexts were present concurrently than if either context was present alone. The temporal network provides a mechanism for altering the available reasoning methods based on the set of current contexts. In this use of the temporal network, reasoning methods are associated with each context. When a context is present, a temporal network node representing that context is created and the reasoning methods are made available to the interpretation process. A temporal network node may also withdraw methods made available by other temporal network nodes. In this manner, a general rule or method can be suspended if it is not appropriate in particular context.

We believe that the combination of mathematical models along with specialized symbolic structures results in more representational and inferencing power than either method alone. Well established mathematical techniques convert observations into underlying system concepts while symbolic techniques interpret the mathematical results using additional domain knowledge. Although some of these features could possibly be represented using either mathematical or symbolic techniques alone, we