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.

*E. Funding Support Status*

1) Representation and Use of Causal Knowledge for Inference from Databases
Robert L. Blum, M.D., Ph.D.: Principal Investigator
National Science Foundation: IST 83-17858
Total award: $89,597 (direct + indirect)
Term: March 15, 1984 through March 14, 1986

2) Deriving Knowledge from Clinical Databases
Gio C. M. Wiederhold, Ph.D.: Principal Investigator
National Library of Medicine: LM-04334
Total award: $291,192 (direct)
Term: May 1, 1984 through November 30, 1986

## 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.

## 6.2. National AIM Projects

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

# 6.2.1. CADUCEUS Project

CADUCEUS Project

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

## I. SUMMARY OF RESEARCH PROGRAM

*A. Project rationale*

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

A major achievement of this research undertaking has been the design of a program called INTERNIST-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.


*E. Funding support*

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

   5 R24 RR01101-08
   07/01/80 - 06/30/85 - $1,607,717
   07/01/84 - 06/30/85 - $354,211

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

   5 R01 LM03710-05
   07/01/80 - 06/30/85 - $817,884
   07/01/84 - 06/30/85 - $210,091

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

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

## II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

### *A, B. Medical Collaborations and Program Dissemination Via SUMEX*

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

### *C. Critique of Resource Management*

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

## III. RESEARCH PLANS

### *A. Project Goals and Plans*

Continued effort to complete the medical knowledge 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.

# 6.2.2. CLIPR - Hierarchical Models of Human Cognition

**Hierarchical Models of Human Cognition (CLIPR Project)**

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

## I. SUMMARY OF RESEARCH PROGRAM

### *A. Project Rationale*

The two CLIPR projects have made progress during the last year. The prose comprehension project has completed one major project, and is designing a prose comprehension model that reflects state-of-the-art knowledge from psychology (van Dijk & Kintsch, 1983) and artificial intelligence. During the last 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.

*F. Funding Support Status*

1. Text Comprehension and Memory
   Walter Kintsch, Professor, University of Colorado
   National Institute of Mental Health - 5 R01 MH15872-14-16
   7/1/84 - 6/30/87: $145,500 (direct)
   7/1/83 - 6/30/84: $56,501

2. Understanding and solving word arithmetic problems
   Walter Kintsch, Professor, University of Colorado
   National Science Foundation
   8/1/83 - 7/31/86: $200,000

3. The Application of Cognitive Complexity Theory to
   the Design of User Interface Architectures
   David Kieras, Associate Professor, University of Michigan

Peter G. Polson, Professor, University of Colorado
International Business Machines Corporation
1/1/85 - 12/31/85: $250,000 (direct+indirect)

## 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.

# 6.2.3. 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.

*E. Funding Support*

Title: MENTOR: Monitoring Drug Therapy for Hospitalized Patients

Principal Investigators:
Terrence F. Blaschke, M.D.
Division of Clinical Pharmacology
Department of Medicine
Stanford University

Stuart M. Speedie, Ph.D.
School of Pharmacy
University of Maryland

Funding Agency: National Center for Health Services Research

Grant Identification Number: 1 R18 HS05263

Total Award: January 1, 1985 - December 31, 1988 485,134 Total Direct Costs

Current Period: January 1, 1985 - December 31, 1985 147,170 Total Direct Costs

## II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

### *A. Medical Collaborations and Program Dissemination via SUMEX*

This project represents a collaboration between faculty at Stanford University Medical Center 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.

# 6.2.4. 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].