used within the system to represent uncertainty, called evoking strengths and frequencies, are poorly defined. This makes it difficult to tune the method used by the program to assign likelihoods to diseases (the scoring scheme) and makes it difficult to transport knowledge contained in the program to other medical diagnostic systems. We have carried out several experiments involving the use of probability theory to better characterize the quantities. These experiments have been performed in collaboration with R. Miller and D. Heckerman.

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

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

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

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

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

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

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

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

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

C.2 Research in Progress

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

D. Publications

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

26. Kuhn, I., Wiederhold, G., Rodnick, J.E., Ramsey-Klee, D.M., Benett, S., Beck, D.D.: *Automated Ambulatory Medical Record Systems in the U.S.*, to be published by Springer-Verlag, 1983, in *Information Systems for Patient Care*, B. Blum (ed.), Section III, Chapter 14.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

E. Funding Support Status

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

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations

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

B. Interactions with Other SUMEX-AIM Projects

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

C. Critique of Resource Management

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

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

III. RESEARCH PLANS

A. Project Goals and Plans

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

SHORT-TERM GOALS --

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

LONG-RANGE GOALS --

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

B. Justification and Requirements for Continued Use of SUMEX

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

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

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

C. Recommendations for Resource Development

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

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

Responses to Questions Regarding Resource Future

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

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

Software for community work stations. Yes. Making good utility programs available to all users sounds like a good idea.

Q: Will you require continued access to the SUMEX-AIM 2060 and if so, for how long?

A: Yes. For the foreseeable future and for the above reasons.

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

A: We would pay them. The 2060 is worth it to us. Of course, if the fees were high, we would consider alternatives.

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

A: We are currently using two of the SUMEX Xerox 1108's for the development of our project. We will stay with these for the foreseeable future.

IV.B. National AIM Projects

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

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

IV.B.1. INTERNIST-I Project

CADUCEUS Project (INTERNIST-I)

This project is unfunded at the present time.

J. D. Myers, M.D.
University Professor Emeritus (Medicine)
University of Pittsburgh
1291 Scaife Hall
Pittsburgh, Pa., 15261

I. SUMMARY OF RESEARCH PROGRAM

A. Project rationale

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

A major achievement of this research undertaking has been the design of a program called INTERNIST-1, along with an extensive medical knowledge base. This program has been used over the past decade to analyze many hundreds of difficult diagnostic problems in the field of internal medicine. These problem cases have included cases published in medical journals (particularly Case Records of the Massachusetts General Hospital, in the New England Journal of Medicine), CPCs, and unusual problems of patients in our Medical Center. In most instances, but by no means all, INTERNIST-I has performed at the level of the skilled internist, but the experience has highlighted several areas for improvement.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical relevance.

The development of the QUICK MEDICAL REFERENCE (QMR) under the leadership of Dr. Randolph A. Miller has allowed us to distribute the INTERNIST-I knowledge base in a modified format to over twenty other academic medical institutions. The knowledge base can thereby be used as an "electronic textbook" in medical education at all levels -- by medical students, residents and fellows, and faculty and staff physicians. This distribution is continuing to expand.

The INTERNIST-I program has been used in recent years to develop patient management problems for the American College of Physician's Medical Knowledge Self-assessment Program, and to develop patient management problems and test cases for the Part III Examination and the developing computerized testing program of the National Board of Medical Examiners.

C. Highlights of Research Progress

C.1 Accomplishments this past year

For the record, it should be noted that grant support for the QMR project has come solely from the CAMDAT Foundation of Farmington, Conn., from the Department of Medicine of the University of Pittsburgh, and from Dr. Miller's NLM RCDA grant. The NLM and DRR grants currently supporting the CADUCEUS project do not in any way support the QMR project.

The group of us (Myers, Miller and Masarie) together with assigned residents in internal medicine and fellows in medical informatics are continuing to expand the knowledge base and to incorporate the diagnostic consultative program into QMR. The computer program for the interrogative part of the diagnostic program is the main remaining task. An editor for the QMR knowledge base, as modified from the INTERNIST-I knowledge base, has been written from scratch in Turbo Pascal by Dr. Masarie. The entire QMR program can be accommodated in, maintained (particularly edited) and operated on individual IBM PC-AT computers.

Our group has incorporated into the QMR diagnostic consultant program modifications and embellishments of the INTERNIST-I knowledge base, and will continue to do so over the next year by adding "facets" of diseases or syndromes. This addition and modification is expected to improve the performance of the diagnostic consultant program.

The medical knowledge base has continued to grow both in the incorporation of new diseases and the modification of diseases already profiled so as to include recent advances in medical knowledge. Several dozen new diseases have been profiled during the past year. The current number of diseases in the QMR knowledge base is 577, and over 4100 possible patient findings are included.

C.2 Research in progress

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

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

2. Institution of field trials of QMR on the clinical services in internal medicine at the Health Center of the University of Pittsburgh. This has been accomplished in a limited fashion beginning April 1987; a "computer-based diagnostic consultation service" has been made available to attending physicians and housestaff on the medical services of our two main teaching hospitals. Institutional Review Board (IRB) approval was granted to the service before it was initiated.

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

4. Adaptation of the diagnostic program and data base of INTERNIST-I and the QMR modification to subserve educational purposes and the evaluation of clinical performance and competence.

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

D. List of relevant publications

1. Myers, J.D.: Educating future physicians: Something old, Something new. Ohio State Univ. Proceedings of Symposium, Medical Education in the 21st Century. 1985.

2. Myers, J.D.: The process of clinical diagnosis and its adaptation to the computer. In The Logic of Discovery and Diagnosis in Medicine, University of Pittsburgh Series in the Philosophy and History of Science, edited by Kenneth F. Schaffner, Univ. of California Press, pp. 155-180, 1985.

3. Masarie, Jr. F.E., Miller, R.A., First, M.B., Myers, J.D.: An Electronic Textbook of Medicine. Proceedings of Ninth Annual Symposium on Computer Applications in Medical Care. Baltimore, Maryland, November 1985.

4. Masarie, Jr. F.E., Myers, J.D., Miller, R.A.: INTERNIST-I PROPERTIES: Representing Common Sense on Good Medical Practice in a Computerized Medical Knowledge Base. Computers and Biomedical Research. 18: 458-479, October 1985.

5. Myers, J.D., Chairman. Medical Education in the Information Age. Proceedings of the Symposium on Medical Informatics. Association of American Medical Colleges, 1986.

6. Miller, R.A., Schaffner K.F., Meisel, A. Ethical and Legal Issues Related to the Use of Computer Programs in Clinical Medicine. Annals of Internal Medicine. 1985; 102:529-36.

7. Miller, R.A., Masarie, F.E., Myers J.D. "Quick Medical Reference" for diagnostic assistance. MD Computing. 1986; 3:34-48.

8. Miller, R.A., Mc Neil, M.A., Challinor, S., Masarie, F.E., Myers, J.D. Status Report: The INTERNIST-1/Quick Medical Reference Project. West J Med. 1986; 145:816-22.

9. Miller, R.A. From Automated Medical Records to Expert System Knowledge Bases: Common Problems in Representing and Processing Patient Data. Topics in Health Record Management, March 1987, in press.

10. Masarie, F.E., Miller, R.A. Medical Subject Headings and Medical Terminology: An analysis of terminology used in hospital charts. Bulletin of the Medical Library Association, 1987; 75:89-94.

11. Miller, R.A., Masarie, F.E., Miller, R.A. Quick Medical Reference (QMR): A microcomputer-based adaptation of the INTERNIST-1 diagnostic system for general internal medicine. In: R. Salamon, B. Blum, M. Jorgenson (eds), MEDINFO 86, p. 1143. Amsterdam: North Holland Publishing Co, 1986.

12. Heckerman, D., Miller, R.A. Towards a better understanding of the INTERNIST-1 knowledge base. In: R. Salamon, B. Blum, M. Jorgenson (eds), MEDINFO 86, pp. 22-26. Amsterdam: North Holland Publishing Co, 1986.

13. McNeil, M.A., Challinor, S.M., Miller, R.A. Preliminary Evaluation of a computer-based medical decision support system in the clinical setting. In Press, Medical Decision-Making, December 1986.

E. Funding support

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

   5 R24 RR01101-08
   07/01/80 - 03/31/86 - $1,658,347
   07/01/84 - 09/30/85 - $354,211
   09/30/85 - 03/31/86 - $50,690

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

   5 R01 LM03710-05
   07/01/80 - 03/31/86 - $853,200
   07/01/84 - 09/30/85 - $210,091
   09/30/85 - 03/31/86 - $35,316

3. Diagnostic-Internist: A Computerized Medical Consultant
   Randolph A. Miller, M.D.
   Associate Professor of Medicine
   University of Pittsburgh Department of Medicine
   National Library of Medicine - Development Award Research
   Career
   National Institutes of Health

   1 KO4 LM00084-01
   09/30/85 - 09/29/90 - amounts to be determined annually
   09/30/85 - 09/29/86 - $55,296
   09/30/86 - 09/29/87 - $55,296

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A,B. Medical Collaborations and Program Dissemination Via SUMEX

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

C. Critique of Resource Management

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

III. RESEARCH PLANS

A. Project Goals and Plans

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

B. Justification and Requirements for Continued SUMEX Use

Our use of SUMEX will obviously decline with the adaptation of our programs to the IBM PC-AT. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended for the present to keep INTERNIST-1 at SUMEX for comparative use as QMR is developed here.

Our best prediction is that our project will require continued access to the 2060 for the next year or two and we consider such access essential to the future development of our knowledge base. After that time, our work can probably be accomplished on our personal work stations.

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

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

IV.B.2. CLIPR - Hierarchical Models of Human Cognition

Hierarchical Models of Human Cognition (CLIPR Project)

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

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The two CLIPR projects have made progress during the last year. The prose comprehension project has completed one major project, and is designing a prose comprehension model that reflects state-of-the-art knowledge from psychology (van Dijk & Kintsch, 1983) and artificial intelligence. During the last five years, Polson, in collaboration with Dr. David Kieras of the University of Michigan, has continued work on a project studying the psychological factors underlying device complexity and the difficulties that nontechnically trained individuals have in learning to use devices like word processors. They have developed formal representations of a user's knowledge of how to operate a device and of the user-device interface (Kieras & Polson, 1985) and have completed several experiments evaluating their theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986).

Technical Goals

The CLIPR project consists of two subprojects. The first, the text comprehension project, is headed by Walter Kintsch and is a continuation of work on understanding of connected discourse that has been underway in Kintsch's laboratory for several years. The second, the device complexity project, is headed by Peter Polson in collaboration with David Kieras of the University of Michigan. They are studying the learning and problem solving processes involved in the utilization of devices like word processors or complex computer controlled medical instruments (Kieras & Polson, 1985).

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension. We hope to be able to merge the substantial artificial intelligence research on these systems with psychological interpretations of prose comprehension, resulting in a computational model that is also psychologically respectable.

The goal of the device complexity project is to develop explicit models of the user-device interaction. They model the device as a nested automata and the user as a production system. These models make explicit kinds of knowledge that are required to operate different kinds of devices and the processing loads imposed by different implementations of a device.

B. Medical Relevance and Collaboration

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

The device complexity project has two primary goals: the development of a cognitive theory of user-device interaction in including learning and performance models, and the development of a theoretically driven design process that will optimize the relationships between device functionality and ease of learning and other performance factors (Polson & Kieras, 1983, 1984; Polson, Muncher, and Engelbeck 1985). The results of this project should be directly relevant to the design of complex, computer controlled medical equipment. They are currently using word processors to study user-device interactions, but principles underlying use of such devices should generalize to medical equipment.

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

C. Highlights of Research Progress

The version of the prose comprehension model of 1978 (Kintsch & van Dijk, 1978), which originally was realized as a computer simulation by Miller & Kintsch (1980), has been extended in a major simulation program by Young (1985). Unlike the earlier program, Young includes macroprocessing in her model, and thereby greatly extends the usefulness of the program. It is expected that this program will be widely useful in studies of prose where a detailed theoretical analysis is desired.

The general theory has been reformulated and expanded in van Dijk & Kintsch (1983). This research report of book length presents a general framework for a comprehensive theory of discourse processing. It has been applied to an interesting special case, the question of how children understand and solve word arithmetic problems, by Kintsch & Greeno (1985). A simulation for this model, using INTERLISP, has been supplied in Fletcher (1985).

The device complexity project is in its fifth year. They have developed an explicit model for the knowledge structures involved in the user-device interaction, and they are developing simulation programs. Their preliminary theoretical results are described in Kieras & Polson (1985). They have also completed several experiments evaluating the theory (Polson & Kieras, 1984, 1985; Polson, Muncher, and Engelbeck, 1986) and have shown that number of productions predicts learning time and that number of cycles and working memory operations predicts execution time for a method.

D. List of Relevant Publications

1. Fletcher, R.C.: *Understanding and solving word arithmetic problems: A computer simulation*. Technical Report No. 135, Institute of Cognitive Science, Colorado, 1984.

2. Kieras, D.E. and Polson, P.G.: *The formal analysis of user complexity*. Int. J. Man-Machine Studies, 22, 365-394, 1985.

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

4. Kintsch, W. and Greeno, J.G.: *Understanding and solving word arithmetic problems.* Psychological Review, 1985, 92, 109-129.

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

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

7. Polson, P.G. and Kieras, D.E.: *A formal description of users' knowledge of how to operate a device and user complexity.* Behavior Research Methods, Instrumentation, & Computers, 1984, 16, 249-255.

8. Polson, P.G. and Kieras, D.E.: *A quantitative model of the learning and performance of text editing knowledge.* In Borman, L. and Curtis, B. (Eds.) Proceedings of the CHI 1985 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 207-212, 1985.

9. Polson, P.G. and Jeffries, R.: *Instruction in general problem solving skills: An analysis of four approaches.* In (Eds.) Siegel, J., Chipman, S., and Glaser, R. Thinking and learning skills: Relating instructions to basic research: Vol. 1. Hillsdale, N.J.: OpLawrence Erlbaum Associates, pp. 414-455.

10. Polson, P.G., Muncher, E., and Engelbeck, G.: *Test of a common elements theory of transfer.* In Mantei, M. and Orbeton, P. (Eds.) Proceedings of the CHI 1986 Conference on Human Factors in Computing. New York: Association for Computing Machinery. pp. 78-83, 1986.

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

12. Young, S.: *A theory and simulation of macrostructure.* Technical Report No. 134, Institute of Cognitive Science, Colorado, 1984.

13. Walker, H.W., Kintsch, W.: *Automatic and strategic aspects of knowledge retrieval.* Cognitive Science, 1985, 9, 261-283.

E. Funding Support

1. Text Comprehension and Memory
   Walter Kintsch, Professor, University of Colorado
   National Institute of Mental Health - 5 R01 MH15872-14-16
   7/1/84 - 6/30/87: $197,500 (direct)

2. Understanding and solving word arithmetic problems
   Walter Kintsch, Professor, University of Colorado
   National Science Foundation
   8/1/83 - 7/31/86: $200,000
   8/1/86 - 7/31/87: $55,400

3. Theories, Methods, and Tools for the Design of User-centered Computer Systems
   Walter Kintsch, Professor, University of Colorado
   Gerhard Fischer, Assoc. Prof. University of Colorado
   Army Research Institute
   8/1/86 - 7/31/91: $500,000
   8/1/86 - 7/31/87: $86,500

4. Software Design for a Propositionalizer
   Walter Kintsch, Professor, University of Colorado
   A. Turner, Research Assoc., University of Colorado
   Air Force Office of Scientific Research
   10/1/85 - 9/30/87: $110,000
   10/1/86 - 9/30/87: $47,000

4. The Application of Cognitive Complexity Theory to the Design of User Interface Architectures
   David Kieras, Associate Professor, University of Michigan
   Peter G. Polson, Professor, University of Colorado
   International Business Machines Corporation
   1/1/85 - 4/31/87: $500,000 (direct+indirect)
   1/1/86 - 4/31/87: $250,000 (direct+indirect)

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

B. Sharing and Interactions with Other SUMEX-AIM Projects

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

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

C. Critique of Resource Management

We have found the staff of SUMEX to be cooperative and effective in dealing with special requirements and in responding to our questions. The facilities for communication on the ARPANET have also facilitated collaborative work with investigators throughout the country.

III. RESEARCH PLANS

A. Long Range Projects Goals and Plans

The goal of the prose comprehension project is to develop a computer system capable of the meaningful processing of prose. This work has been generally guided by the prose comprehension model discussed by van Dijk & Kintsch (1983), although our programming efforts have identified necessary clarifications and modifications in that model (Kintsch & Greeno, 1985; Fletcher, 1985; Walker & Kintsch, 1985; Young, 1985). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension. We hope to be able to merge the substantial artificial intelligence research on these systems with psychological interpretations of prose comprehension, resulting in a computational model that is also psychologically respectable.

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

B. Justification and Requirements for Continued SUMEX Use

Both the prose comprehension and the user-computer interaction projects have shifted their actual simulation work from SUMEX to systems at the University of Colorado and the University of Michigan. Both projects use Xerox 1108 systems continuing their work in INTERLISP.

Access to SUMEX's mail facilities are critical for the continued success of these projects. These facilities provide us with the means to interact with colleagues at other universities. Kintsch is currently collaborating with James Greeno, who is at the University of California at Berkeley, and Polson's long-term collaborator, David Kieras, is at the University of Michigan. In addition, our access to the Xerox 1108 (Dandelion) user's community is through SUMEX.

We currently use five computing systems: a VAX 11/780, a MicroVAX II, and three Xerox 1108s, one of which is at the University of Michigan. The VAX's are used to collect experimental data designed to evaluate the simulation models and to do necessary statistical analysis.

C. Needs and Plans for Other Computational Resources

SUMEX provides us with communication which we discussed in the preceding paragraph.

D. Recommendations for Future Community and Resource Development

We will continue to need access to the SUMEX-AIM 2060 in order to access communication networks.

IV.B.3. MENTOR Project

MENTOR Project

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

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

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

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

In order to detect and appropriately manage adverse drug reactions, sophisticated medical knowledge and problem solving is required. Expert systems offer the possibility of embedding this expertise in a computer system. Such a system could automatically gather the appropriate information from existing record-keeping systems and continually monitor for the occurrence of adverse drug reactions. Based on a knowledge base of relevant data, it could analyze incoming data and inform physicians when adverse reactions are likely to occur or when they have occurred. The MENTOR project is an attempt to explore the problems associated with the development and implementation of such a system and to implement a prototype of a drug monitoring system in a hospital setting.

B. Medical Relevance and Collaboration

A number of independent studies have confirmed that the incidence of adverse reactions to drugs in hospitalized patients is significant and that they are for the most part preventable. Moreover, such statistics do not include instances of suboptimal drug therapy which may result in increased costs, extended length-of-stay, or ineffective therapy. Data in these areas are sparse, though medical care evaluations carried out as part of hospital quality assurance programs suggest that suboptimal therapy is common.

Other computer systems have been developed to influence physician decision making by monitoring patient data and providing feedback. However, most of these systems suffer from a significant structural shortcoming. This shortcoming involves the evaluation rules that are used to generate feedback. In all cases, these criteria consist of discrete, independent rules, yet medical decision making is a complex process in which many factors are interrelated. Thus, attempting to represent medical decision-making as a discrete set of independent rules, no matter how complex, is a task that can, at best, result in a first-order approximation of the process. This places an inherent limitation on the quality of feedback that can be provided. As a consequence it is extremely difficult to develop feedback that explicitly takes into account all information available on the patient. One might speculate that the lack of widespread acceptance of such systems may be due to the fact that their recommendations are often rejected by physicians. These systems must be made more valid if they are to enjoy widespread acceptance among physicians.

The proposed MENTOR system is designed to address the significant problem of adverse drug reactions by means of a computer-based monitoring and feedback system to influence physician decision-making. It will employ principles of artificial intelligence to create a more valid system for evaluating therapeutic decision-making.

The work in the MENTOR project is a collaboration between Dr. Blaschke at Stanford University, Dr. Speedie at the University of Maryland, and Dr. Charles Friedman at the University of North Carolina. Dr. Speedie provides the expertise in the area of artificial intelligence programming. Dr. Blaschke provides the medical expertise. Dr. Friedman contributes expertise in the area of physician feedback design and system impact evaluation. The blend of previous experience, medical knowledge, computer science knowledge and evaluation design expertise they represent is vital to the successful completion of the activities in the MENTOR project.

C. Highlights of Research Progress

The MENTOR project was initiated in December, 1983. The project has been funded by the National Center for Health Services Research since January 1, 1985. Initial effort focused on exploration of the problem of designing the MENTOR system. As of June 1, 1987, a working prototype system has been developed and is undergoing evaluation. The prototype consists of a Patient Data Base, an Inference Engine, an Advisory Module and a Medical Knowledge Base. The Medical Knowledge Base currently contains information related to Aminoglycoside Therapy, Digoxin therapy, Surgical Prophylaxis, and Microbiology Lab reports. The system is currently implemented on a Xerox 1186 AI Workstation. Another version of the Patient Data Base has been developed for a VAX 750 and is currently being tested. Plans call for the interconnection of the VAX and the 1186 running the inference engine. The VAX will then be connected to a Hospital Information System for data acquisition.

E. Funding Support

Title: MENTOR: Monitoring Drug Therapy for Hospitalized Patients

Principal Investigators:

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

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

Funding Agency: National Center for Health Services Research

Grant Identification Number: 1 R18 HS05263

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

Current Period: January 1, 1987 - December 31, 1987 $195,731 Total Direct Costs

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Dissemination via SUMEX

This project represents a collaboration between faculty at Stanford University Medical Center, the University of Maryland School of Pharmacy, and the University of North Carolina in exploring computer-based monitoring of drug therapy. SUMEX, through its communications capabilities, facilitates this collaboration of geographically separated project participants by providing electronic mail and file exchange between sites.

B. Sharing and Interactions with Other SUMEX-AIM Projects

Interactions with other SUMEX-AIM projects has been on an informal basis. Personal contacts have been made with individuals working on the ONCOCIN project concerning system development issues. Dr. Perry Miller has also been of assistance by providing software for advisory generation. Given the geographic separation of the investigators, the ability to exchange mail and programs via the SUMEX system as well as communicate with other SUMEX-AIM projects is vital to the success of the project.

C. Critique of Resource Management

To date, the resources of SUMEX have been fully adequate for the needs of this project. The staff have been most helpful with any problems we have had and we are quite satisfied with the current resource management.

III. RESEARCH PLANS

A. Project Goals and Plans

The MENTOR project has the following goals:

1. Implement a prototype computer system to continuously monitor patient drug therapy in a hospital setting. This will be an expert system that will use a modular, frame-oriented form of medical knowledge, a separate inference engine for applying the knowledge to specific situations, and automated collection of data from hospital information systems to produce therapeutic advisories.

2. Select a small number of important and frequently occurring medical settings (e.g., combination therapy with cardiac glycosides and diuretics) that can lead to therapeutic misadventures, construct a comprehensive medical knowledge base necessary to detect these situations using the information typically found in a computerized hospital information system and generate timely advisories intended to alter behavior and avoid preventable drug reactions.

3. Design and begin to implement an evaluation of the impact of the prototype MENTOR system on physicians' therapeutic decision-making as well as on outcome measures related to patient health and costs of care.

1987 will be spent on continued prototype development in four content areas, refinement of the inference mechanisms, and interfacing to existing patient information systems.

B. Justification and Requirements for Continued SUMEX Use

This project needs continued use of the SUMEX facilities for two reasons. First, it provides access to an environment specifically designed for the development of AI systems. The MENTOR project focuses on the development of such a system for drug monitoring that will explore some neglected aspects of AI in medicine. This environment is necessary for the timely development of a well-designed and efficient MENTOR system. Second, access to SUMEX is necessary to support the collaborative efforts of geographically separated development teams at Stanford and the University of Maryland.

Furthermore, the MENTOR project is predicated on the access to the SUMEX resource free of charge over the next two years. Given the current restrictions on funding, the scope of the project would have to be greatly reduced if there were charges for use of SUMEX.

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

A major long-range goal of the MENTOR project is to implement this system on a independent hardware system of suitable architecture. It is recognized that the full monitoring system will require a large patient data base as well as a sizeable medical knowledge base and must operate on a close to real-time basis. Ultimately, the SUMEX facilities will not be suitable for these applications. Thus, we have transported the prototype system to a dedicated hardware system that can fully support the the planned system and which can be integrated into a Hospital Information System. For this purpose a VAX 750 and three Xerox 1186 workstations have been acquired and our development efforts have been transferred to them.

D. Recommendations for Future Community and Resource Development

In the brief time we have been associated with SUMEX, we have been generally pleased with the facilities and services. However, it is clearly evident that the users' almost insatiable demands for CPU cycles and disk space cannot be met by a single central machine. The best strategy would appear to be one of emphasizing powerful workstations or relatively small, multi-user machines linked together in a nationwide network with SUMEX serving as the its central hub. This would give the individual users much more control over the resources available for their needs, yet at the same time allow for the communications among users that have been one of SUMEX's strong points.

For such a network to be successful, further work needs to be done in improving the network capabilities of SUMEX to encourage users at sites other than Stanford. Further work is also needed in the area of personal workstations to link them to such a network. Given the successful completion of this work, it would be reasonable to consider the gradual phase-out of the central SUMEX machine over two or three years and its replacement by an efficient, high-speed communications server.

IV.B.4. SOLVER Project

SOLVER: Problem Solving Expertise

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

Dr. James R. Slagle
Department of Computer Science
University of Minnesota

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

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

The SOLVER project is an interdisciplinary research effort concerned with understanding medical expertise, particularly in diagnostic tasks. The Minnesota SOLVER project focuses upon the development of strategies for discovering and representing the knowledge and skill of expert problem solvers. Although in the last fifteen years considerable progress has been made in synthesizing the expertise required for solving complex problems, most expert systems embody only a limited amount of expertise. What is still lacking is a theoretical framework capable of reducing dependence upon the expert's intuition or on the near-exhaustive testing of possible organizations. Our methodology consists of: (1) extensive use of verbal thinking aloud protocols as a source of information from which to make inferences about underlying knowledge structures and processes; (2) development of computer models as a means of testing the adequacy of inferences derived from protocol studies; (3) testing and refinement of the cognitive models based upon the study of human and model performance in experimental settings. Currently, we are investigating problem-solving expertise in domains of medicine, financial auditing, management, and law.

B. Medical Relevance and Collaboration

Much of our research has been and will continue to be directly focused on medical AI problems. Medical diagnostic expertise is a complex phenomenon which is not yet fully understood. The SOLVER project is studying both the theoretical foundations of expertise and also is engaged in the design and testing of medical expert systems.

A medical expert system in pediatric cardiology has been designed in collaboration with Dr. James Moller, Department of Pediatrics, University of Minnesota Hospitals.

Dr. Donald Connelly, Department of Laboratory Medicine, University of Minnesota School of Medicine, has supervised a number of medical expert system projects, including projects in analysis of time series observations and platelet transfusion practice.

Dr. Slagle's research group has developed an expert system shell called AGNESS ("A Generalized Network-based Expert System Shell"), and has developed three medical expert systems that either use AGNESS or are modeled after AGNESS. AGNESS uses a computation network rather than a production rule base and supports values of any well-defined data type, the Merit questioning scheme, an explanation facility, and expert-defined inference methods. The first major application of AGNESS was to implement the clinical expert system ETA (Exercise Test Analyzer). The cases studied came from the Program on the Surgical Control of the Hyperlipidemas (POSCH), a study of the effect of reduced cholesterol on heart attack victims.

Kent Spackman, M.D. was a post-doctoral fellow in medical informatics at the University of Minnesota who is completing a Ph.D. thesis in Artificial Intelligence at the University of Illinois. During his residency at the University of Minnesota Hospitals, Dr. Spackman collaborated with the SOLVER project. Dr. Spackman's research addressed issues in automated knowledge acquisition for medical expert systems.

C. Highlights of Research Progress

Accomplishments of This Past Year

Dr. Connelly has continued supervising the development of an expert system, ESPRE, to be used in monitoring requests for platelet transfusions. The prototype knowledge base was refined and extended, communications protocols to communicate with laboratory computer systems have been improved, a standing order feature has been implemented, the inference engine has been modified, and a preliminary evaluation has been completed. In the evaluation, 68 transfusion requests were processed by the system. In more than 80% of the cases, the expert system agreed with the blood bank decision to transfuse platelets. In six of the remaining cases, the expert system declined to propose a decision because there was no recent platelet count available to it. In four cases, additional clinical factors known to the blood bank physician were brought to bear, and the transfusions were authorized even though usual transfusion criteria had not been met. The expert system is being placed in parallel operation in the blood bank to be used as a consultation tool.

In addition, Dr. Connelly is supervising the project dealing with detection of deviations in time series by the human observer. This project involves the implementation of a number of small expert systems used in modeling the human graph reader. During the past year the work has been extended by examining individual observer differences in deviation detection performance and approach to graph reading. Time trend graphs representing monthly monitoring of serum carcinoembryonic antigen (CEA) levels in simulated patients with surgically-removed breast cancer were presented to twelve clinical laboratory observers and a time series analysis (TSA) routine which is based on a homeostatic model. The observers described their rationale in assigning a level of suspicion regarding the presence of an important deviation as the observation points were serially revealed to them. The verbalization reports were analyzed to develop rules that described each reader's graph reading strategy. Strategies were compared for commonality and difference. Rules obtained from the top two observers were merged into a common rule base and an expert system implemented. A second expert system was constructed by merging consistent rules from all observers. The deviation detection performance of all three approaches (TSA, observers, and both expert systems) are being compared. The analysis is currently in progress.

Dr. Johnson's research group has developed an expert system inference engine called "Cleric." Cleric is a rule based language written in Common Lisp which resembles a forward chaining production system. Cleric has been written to investigate diagnostic problem solving tasks. Cleric differs from simple production systems because it can dynamically create new specialized forms of existing rules for later execution. In addition, Cleric uses a subset of de Kleer's assumption based truth maintenance system (ATMS). A computer hardware diagnosis expert system called "Vesalius" has been implemented in Cleric.