Another class of learning mechanisms are concerned with tuning existing procedures so that they apply more appropriately. There are various mechanisms concerned with extending or generalizing the range of application of a procedure. In the past year we have been working at reducing these different generalization processes to a common partial matching process. In addition to generalization, tuning occurs in the ACTF system by means of discrimination and composition. Discrimination is a process for restricting the range of applicability of a production. Composition attempts to build macro-operators out of a series of productions.

The third direction of our learning work has been concerned with developing a flexible strength-based set of conflict resolution rules. Here we are concerned with modelling the gradual improvement seen in human cognitive skills and also providing the system with the resilience so that it can recover from noise and changes in environmental contingencies.

We have been applying this theory in detail to a simulation of how students acquire proof skills in geometry. We have a more or less thorough analysis of how students learn new postulates of geometry; we initially use these postulates in an interpretative fashion, integrating them with prior knowledge; how they compile special purpose procedures that directly apply this knowledge to proof generation; and how these procedures become tuned with practice. This application has provided strong evidence for most of the learning developments in the ACT system. It has also forced us to develop formalisms for how planning and problem-solving should be structured within a production-system framework.

D. List of Project Publications

[1] Anderson, J.R. Language, Memory, and Thought. Hillsdale, N.J.: L. Erlbaum, Assoc., 1976.

[2] Kline, P.J. & Anderson, J.R. The ACTE User's Manual, 1976.

[3] Anderson, J.R., Kline, P. & Lewis, C. Language processing by production systems. In P. Carpenter and M. Just (Eds.). Cognitive Processes in Comprehension. L. Erlbaum Assoc., 1977.

[4] Anderson, J.R. Induction of augmented transition networks. Cognitive Science, 1977, 125-157.

[5] Anderson, J.R. & Kline, P. Design of a production system. Paper presented at the Workshop on Pattern-Directed Inference Systems, Hawaii, May 23-27, 1977.

[6] Anderson, J.R. Computer simulation of a language acquisition system: A second report. In D. LaBerge and S.J. Samuels (Eds.). Perception and Comprehension. Hillsdale, N.J.: L. Erlbaum Assoc., 1978.

[7] Anderson, J.R., Kline, P.J., & Beasley, C.M. A theory of the acquisition of cognitive skills. In G.H. Bower (Ed.). Learning and Motivation, Vol. 13. New York: Academic Press, 1979.

[8] Anderson, J.R., Kline, P.J., & Beasley, C.M. Complex Learning. In R. Snow, P.A. Frederico, & W. Montague (Eds.). Aptitude, Learning, and Instruction: Cognitive Processes Analyses. Hillsdale, N.J.: Lawrence Erlbaum Assoc., 1980.

[9] Anderson, J.R. & Kline, P.J. A learning system and its psychological implications. In the Proceedings of the Sixth International Joint Conference on Artificial Intelligence, 1979.

[10] Reder, L.M. & Anderson, J.R. Use of thematic information to speed search of semantic nets. Proceedings of the Sixth International Joint Conference on Artificial Intelligence, 1979, 708-710.

[11] Neves, D.M. & Anderson, J.R. Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J.R. Anderson (Ed.), Cognitive Skills and their Acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981.

[12] Anderson, J.R., Greeno, J.G., Kline, P.J., & Neves, D.M. Acquisition of Problem-solving skill. In J. R. Anderson (Ed.), Cognitive Skills and their Acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1981.

[13] Anderson,J.R. Tuning of search of the problem space for geometry proofs. Proceedings of IJCAI-81,1981,165-170.

[14] Anderson, J.R. A theory of language acquisition based on general learning principles.Proceedings of IJCAI-81,1981,97-170.

[15] Anderson,J.R. A general learning theory and its application to the acquisition of proof skills in geometry. In R.Michalski, J.Carbonell, and T. Mitchell(Eds) Machine Learning, forthcoming.

[16] Anderson, J.r. A.C.T. Cambridge, MA>: Harvard University press, 1983.

[17] Kline, P. J. The superiority of relative criteria in partial matching and generalization. Proceedings of IJCAI-81, 1981,296-303.

E. Funding Support

An information-processing analysis of learning in geometry
John R. Anderson, Principal Investigator
National Science Foundation (IST-80-15357)
$186,000 Feb 15, 1981 - Feb 15, 1984

II. INTERACTION WITH THE SUMEX-AIM RESOURCE

A. Collaborations, Interactions, and Sharing of Programs via SUMEX.

We have received and answered many inquiries about the ACT system over the ARPANET. This involves sending documentations, papers, and copies of programs. The most extensive collaboration has been with Greeno and Lesgold who were also on SUMEX.

We find the SUMEX-AIM workshops ideal vehicles for updating ourselves on the field and for getting to talk to colleagues about aspects of their work of importance to us.

Due to memory space problems encountered by ACT we expect that soon we will need to make use of the smaller version of INTERLISP developed at SUMEX for use in the CONGEN program.

B. Critique of Resource Management

The SUMEX-AIM resource has been well suited for the needs of our project. We have made the most extensive use of the INTERLISP facilities and the facilities for communication on the ARPANET. We have found the SUMEX personnel extremely helpful both in terms of responding to our immediate emergencies and in providing advice helpful to the long-range progress of the project. Despite the fact that we are not located at Stanford, we have not encountered any serious difficulties in using the SUMEX system; in fact, there are real advantages in being in the Eastern time zone where we can take advantage of the low load on the system during the morning hours. We have been able to get a great deal of work done during these hours and try to save our computer-intensive work for this time.

Two location changes by the ACT project (from Michigan to Yale in the summer of 1976 and from Yale to Carnegie-Mellon in the summer of 1978) have demonstrated another advantage of working on SUMEX: In both cases we were back to work on SUMEX the day after our arrival.

III. RESEARCH PLANS (8/80-7/86)

A. Project Goals and Plans

Our long-range goals are: (1) Continued development of the ACT system; (2) Application of the system to modeling of various cognitive processes; (3) Dissemination of the ACT system to the national AI community.

This is a period of major evolution for the ACT theory. We have been developing three special versions of the ACTF learning that allow us to more efficiently simulate learning in three domains: proving theorems in geometry, speaking a new language, and writing programs in LISP. We are also performing special purpose simulations of the processes of spreading activation in memory retrieval and of pattern-matching processes in reading. We will be assimilating our experiences with these special purpose simulations in putting forth a major revision of the ACT theory. A research monograph is being written setting forth this theory and is scheduled for publication in Spring,1983. Subsequent to the writing of this monograph we intend to implement the ACT* successor to ACTF that will embody the new conceptions.

B. Justification for Continued Use of SUMEX:

Our goal for the ACT system is that it should serve as a ready-made "programming language" available to members of the cognitive science community for assembling psychologically-accurate simulations of a wide range of cognitive processes. Our intention and ability to provide such a resource justifies our use of the SUMEX facility. This facility is designed expressly for the purpose of developing and supporting such national AI resources and is, in this regard, clearly superior to the facilities we have available locally from the Carnegie-Mellon computer science department. Among the most important SUMEX advantages are the availability of INTERLISP on a machine accessible by either the ARPANET or TYMNET and the existence of a GUEST login. It appears that, at least for the time being, ACT has no hope of being a national resource unless it resides at SUMEX.

C. Needs and Plans for Other Computational Resources

Carnegie-Mellon's is upgrading its PDP-10 hardware to emerging state-of-the-art machines (PERQ,VAX, LISP machines, etc.) promises to provide a excellent resource, and we hope to have access to that resource as it develops. A major decision that needs to be made in the next 6 months is whether to implement the ACT* system locally or on SUMEX.

D. Comments and Suggestions for Future Resource Goals

We are beginning to feel squeezed by various limitations of the SUMEX facility. The problem of peak load is quite serious. We have also been struggling with the address limitations of the current INTERLISP which is made more grievous by the amount of space INTERLISP requires. The computation time and address space limitations have meant that we have not been able to pursue certain projects that we would have otherwise. We applaud any efforts to increase computational power, to increase the address space of INTERLISP (e.g. VAXes), or to create significantly more space efficient versions of INTERLISP.

It seems clear that the eventual trend will be to do more and more computation locally on the ever more powerful and cheaper machines. SUMEX could help us in the choices we are facing in terms of local computation. If we could have experimental access to the new machines like Dolphins, related software, and software advise, it would be easier to make decisions about new resources. It would also be very helpful to develop programs on the SUMEX machines which could eventually be transferred to local machines when acquired. Finally, and perhaps most important, it seems important to make an effort to promote communication among local machines to maintain the community that has been built up through SUMEX.

II.A.2.2 CADUCEUS Project (INTERNIST)

CADUCEUS Project

J. D. Myers, M.D. and H. Pople, Ph.D.
University of Pittsburgh
Pittsburgh, Pennsylvania

I. SUMMARY OF RESEARCH PROGRAM

A. Medical 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-I, along with an extensive medical knowledge base. This program has been used over the past 10 years 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 high-lighted several areas for improvement.

The INTERNIST-I system, in complex cases with multiple diagnoses, operates in a sequential manner. The leading diagnostic hypothesis, i.e. the one with the most supportive clinical evidence both positive and negative, is worked on initially until confirmed or downgraded and replaced by a second hypothesis which is now the stronger. When the first diagnosis is confirmed, the system awards credit to all other diseases which are related ("linked") to the confirmed diagnosis so as to favor, as is logical, a set of related diagnoses rather than independent ones. This sequential analysis can be time consuming and at times involves "false starts." The skilled clinician, instead of focusing on disease A, confirming it, and then moving on to related disease B, etc., will consider from the onset of hypothesis of disease A plus disease B, in other words a diagnostic complex. We believe that our second generation system, CADUCEUS, can be built to accomplish this same form of logical analysis. To do so requires refinement of the computer diagnostic algorithms as well as some reorganization of the medical knowledge base.

The concept of "facets" of diseases has been introduced. Facets are selected pathophysiologic processes or concepts which apply to a broad range of individual diseases. They include such items as cholestasis, coagulopathy of liver disease, acute hemorrhage, and respiratory acidosis.

Facets are comprised of individual clinical manifestations of disease, and the number of the latter in a facet varies from a few to as many as fifteen. A more general or higher level facet may include a number of more specific or limited facets in its structure. Drs. Myers and Miller have reorganized the 57 individual hepatobiliary diseases profiled in the medical knowledge base into facets. Some manifestations of disease are specific predisposing factors and others are specific findings, e.g. bacterial cultures, but the vast majority of manifestations are subsumed into facets. It is reassuring that, taking facets, predisposing factors and specific findings together, all significant manifestations of each of the 57 hepatobiliary diseases are included in the reorganization. The number of facets per disease varies from a half-dozen to more than a dozen.

The use of facets provides certain advantages. The diagnostic analysis starts from broader concepts and more basic pathophysiological and patho-biochemical ones. The use of facets should overcome a present disadvantage of INTERNIST-I, namely that the first diagnosis with which a clinical finding is compatible explains that finding even though some other disease present in the patient is a better and more potent explanation. In the new system, individual manifestations are often assigned to multiple facets and thus may be utilized more prudently in the diagnostic analysis.

Manifestations, as well as facets, are being modified so as to portray important anatomical information, a point often underplayed in INTERNIST-I.

It is planned that CADUCEUS will as well include a better time axis in regard to the development of a disease process or the sequence of diseases.

B. Medical Relevance and Collaboration

The program inherently has direct and substantial medical rele-vance.

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 will be expected to aid considerably any future collaboration.

C. Highlights of Research Progress

Accomplishments This Past Year:

During the past year, there has been continued expansion of the CADUCEUS knowledge base within the context of the original INTERNIST-I system, which is operational at SUMEX-AIM. In addition, there has been a concerted effort to develop an in-house capability to enable programs originally written in INTERLISP to be run in our new local VAX/FRANZLISP environment. To this end, we have developed a set of macro capabilities that handle the major dissimilarities of these two LISP host processors. In particular, we now have within FRANZLISP reasonably complete CLISP and RECORD package facilities.

Our objective has been to make operational in the local setting the new CADUCEUS diagnostic programs and knowledge base editors, which had been developed at SUMEX. At this writing, this conversion process is virtually complete. Because of extensive machine dependencies in the original INTERNIST-I code, and because of known deficiencies in the logic and structure of that program there are no current plans to convert INTERNIST-I to run in the VAX/FRANZLISP environment.

The modified manifestations and the newly devised facets are in process of entry into our new VAX computer, and extensive testing of the reorganized medical knowledge base and the computer operating program (in FRANZLISP) is planned for the next year.

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. The knowledge base now in-cludes 536 individual disease profiles, 3,679 manifestations of disease, and about 1,500 "links" or interrelationships among diseases as well as a myriad of miscellaneous pieces of information which are essential for the correct operation of the system. Twenty-two new diseases have been profiled during the past year and the pediatrics knowledge base has continued to grow as well. In addition, neurologists Gordon Banks and Bruce Weimer, who have recently become associated with the project, have begun developing CADUCEUS knowledge structures for selected problem areas in neurology.

Research in Progress:

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

1) The completion, 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 five years' 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. Pople, H.E. "The Formation of Composite Hypotheses in Diag-nostic Problem Solving: An Exercise in Synthetic Reasoning", Proceedings of the Fifth International Joint Conference on Artificial Intelligence, Boston, August 1977.

2. Pople, H.E. "On the Knowledge Acquisition Process in Applied A.I. Systems", Report of Panel on Applications of A.I., Proceedings of Fifth International Joint Conference on Arti-ficial Intelligence, 1977.

3. Pople, H.E., Myers, J. D. & Miller, R.A. "The DIALOG Model of Diagnostic Logic and its Use in Internal Medicine, Proceedings of the Fourth International Joint Conference on Artificial Intelligence, Tbilisi, USSR, September 1975.

4. Pople, H.E. "Artificial Intelligence Approaches to Computer-based Medical Consultation, Proceedings IEEE Intercon, New York, 1975.

5. 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, University of California Press (in press).

6. Myers, J. D., Pople, H. E. & Miller, R. A. "INTERNIST: Can Artificial Intelligence Help?" in Clinical Decision and Laboratory Use, University of Minnesota Press (in press).

7. Pople, H. E. "Coming to Grips with the Multiple Diagnosis Problem," in Computer-Assisted Decision Making Using Clinical and Paraclinical (Laboratory) Data. B. Statland & S. Bauer (eds.) Mediad Inc., Tarrytown, N. Y., 1980, pp. 81-88. Reprinted in The Logic of Discovery and Diagnosis in Medicine, University of Pittsburgh Series in the Philosophy and History of Science, University of California Press (in press).

8. Pople, H. E. "Heuristic Methods for Imposing Structure on Ill-structured Problems: The Structuring of Medical Diagnostics," in Artificial Intelligence in Medicine, AAAS Symposium Series, Westview Press (forthcoming 1982).

E. Funding Support

1. Clinical Decision Systems Research Resource

   Harry E. Pople, Jr., Ph.D.
   Associate Professor Business

   Jack D. Myers, M.D.
   University Professor (Medicine)
   University of Pittsburgh

   Division of Research Resources
   National Institutes of Health

   5 R24 RR01101-05

   07/01/80 - 06/30/85
   $1,607,717

   07/01/81 - 06/30/82
   $202,632

2. INTERNIST: A Computer-Based Diagnostic Consultant

   Harry E. Pople, Jr., Ph.D.
   Associate Professor of Business

   Jack D. Myers, M.D.
   University Professor (Medicine)
   University of Pittsburgh

   National Library of Medicine
   National Institutes of Health

   5 R01 LM03710-02

   07/01/80 - 06/30/85
   $817,884

   07/01/81 - 06/30/82
   $126,746

3. New Computer-Based Patient Case Simulator

Randolph A. Miller, M.D.
Assistant Professor of Medicine
University of Pittsburgh

National Library of Medicine - New Investigator
National Institutes of Health

5 R23 LM03589-02

07/01/80 - 06/30/83
$87,005

07/01/81 - 06/30/82
$29,555

4. Neurologic Consultation Computer Program

Gordon E. Banks, M.D.
Assistant Professor of Medicine
University of Pittsburgh
National Library of Medicine - New Investigator
National Institutes of Health

5 R23 LM03889-01

04/01/82 - 03/31/85
$107,675

04/01/82 - 03/31/83
$35,725

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A, B. Collaborations and Medical Use of Program 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 pur-poses 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 com-munity there continues to be an exchange of information and states of pro-gress. Such interactions particularly take place at the annual AIM Work-shop.

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 (7/81-6/86)

A. Project Goals and Plans

The prototype CADUCEUS programs and the trial reorganization of the liver and biliary tract diseases will be tested in the VAX/FRANZLISP environment over the summer and fall of this year. As rapidly as possible and pending further refinement and reorganization from experience with the new system, the remainder of the medical knowledge base will be entered. Local and later collaborative field trials must necessarily be postponed until this development has been accomplished.

At least 200 important medical diseases remain to be pro-grammed. Renewed effort in this direction is now being expanded now that other tasks have been surmounted. Expanded efforts in the fields of neurology and pediatrics are included as described above.

B. Justification and Requirements for Continued SUMEX Use

Our use of SUMEX will obviously decline with the installation of our VAX. Nevertheless, the excellent facilities of SUMEX are expected to be used for certain developmental work. It is intended, further, to keep INTERNIST-I at SUMEX for comparative use as CADUCEUS is developed here. 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 the dedicated VAX computer recently installed.

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.

II.A.2.3 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 in their research in this past year. This progress is almost completely due to our access to the SUMEX facility. The prose comprehension group has completed one major project, and is currently interacting with other SUMEX projects with the goal of building a prose comprehension model that reflects state-of-the-art knowledge from psychology and artificial intelligence.

The main activity of the planning group during the last year has been completion of the analysis of thinking-out-loud protocols collected from both expert and novice software designers. SUMEX facilities have been used to store, edit, and reformat the raw protocols to facilitate later analysis. Results of successive analyses are then input to SUMEX, and SUMEX facilities are used to collate the various results. In addition, we have just started a new project studying the psychological factors underlying device complexity and the difficulties that nontechnically trained individuals have in learning to use devices like work processors.

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 over eight years. The second, the planning project, is headed by Peter Polson of the University of Colorado and Michael Atwood of Bell Telephone Laboratories,Denver, formerly of Science Applications Incorporated, Denver, and is studying the processes of planning using software design tasks. In addition, Polson in collaboration with David Kieras of the University of Arizona, Tucson, is studying the learning and problem solving processes involved in the utilization of devices like word processors.

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 Kintsch and van Dijk (1978), although our programming efforts have identified necessary clarifications and modifications in that model (Miller & Kintsch, 1980, 1981; Kintsch & Miller, 1981; Miller, 1982). In general, this research has emphasized the importance of knowledge and knowledge-based processes in comprehension, and we are accordingly working with the AGE and UNITS groups at SUMEX toward the development of a knowledge-based, blackboard model of prose 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 planning project is the development of a model of human performance on software design tasks. We have concentrated on developing a deeper understanding of our protocol data, to increase our knowledge of the details of the planning processes and the knowledge structures that experts use in the process of planning. We have developed a method of protocol analysis that essentially involves the transforming of the protocol into a low level theoretical description of the processes used to solve the design problem. We have assumed a very simplified version of a blackboard model that is described in Jeffries, Turner, Polson, and Atwood (1981). We currently carry out our analysis by hand, developing a form of this low level model for each protocol.

The goal of the device complexity project is to develop explicit models of the user-device interaction. Our current plan is to model the device as a nested automata and the user as a production system. The purpose of developing such models would be to make explicit kinds of knowledge that are required to operate different kinds of devices and to specify the processing loads imposed by different implementations of a device. We feel that this offer are tools that are being developed at SUMEX--in particular AGE and the UNIT package--will dramatically facilitate our abilities to generate such models of the user-device interface.

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.

Note that our goal of a blackboard model is particularly relevant to the understanding of learning difficulties. One important aspect of a blackboard model is the separation of cognitive processes into a set of interacting subprocesses. Once such subprocesses have been identified and constructed, it would be instructive to observe the model's performance when certain of these processes are facilitated or inhibited. Many researchers have shown that there are a variety of cognitive deficits (insufficient short-term memory capacity, poor long-term memory retrieval, and such) that can lead to reading problems. Having a blackboard model in which the power of individual components could be manipulated would be a significant step in determining the nature of such reading problems.

The planning project is attempting to gain understanding of the cognitive mechanisms involved in design and planning tasks. The knowledge gained in such research should be directly relevant to a better understanding of the processes involved in medical policy making and in the design of complex experiments. We are currently using the task of software design to describe the processes underlying more general planning mechanisms that are also used in a large number of task oriented environments like policy making.

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

The on-going development of the prose comprehension model would not be possible without our collaboration with the AGE and UNITS research groups. We look forward to a continued collaboration, with, we hope, mutually beneficial results. 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 Pittsburgh. Needless to say, all of this interaction has been greatly facilitated by the local and network-wide communication systems supported by SUMEX. There has been considerable communication between members of the prose comprehension and AGE/UNITS groups as program bugs have been discovered and corrected; the presence of a mail system has made this process infinitely easier than if telephone or surface mail messages were required. 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.

C. Progress Summary

The prose comprehension project has completed an initial version of a model of prose comprehension (Miller & Kintsch, 1980). This model has been applied to a large number of texts, and has yielded quite reasonable predictions of recall and readability. Psychologists from other universities have used this system to derive reading time and recall predictions for their own experimental materials. We are currently using the AGE and UNITS packages to extend this model toward one that can make use of world knowledge in its analyses; this model is discussed in Miller and Kintsch (1981) and Miller (1982).

The planning group has completed the detailed analysis of several long thinking-out-loud protocols collected from both expert and novice software designers. These analyses involved the development of a lower level model for each of the protocols. See Atwood and Jeffries (1980) for details and examples.

The device complexity project is just getting underway. One of our primary objectives for this year is to develop an explicit model for the knowledge structures involved in the user-device interaction and we will then consider designs for possible simulation programs.

D. List of Relevant Publications

Atwood, M. E., & Jeffries, R. Studies in plan construction I: Analysis of an extended protocol. Technical Report SAI-80-028-DEN, Science Applications, Incorporated, Denver, Co. March, 1980.

Atwood, M. E., & Jeffries, R. Studies in plan construction II: Novice design behavior. Technical Report SAI-80-154-DEN, Science Applications, Incorporated, Denver, Colorado, December, 1980.

Atwood, M. E., Polson, P. G., Jeffries, R., and Ramsey, H. R. Planning as a process of synthesis. Technical Report SAI-78-144-DEN, Science Applications, Incorporated, Denver, Co. December, 1978.

Jeffries, R., Turner, A., , Polson, P. G., & Atwood, M. E. The process of designing software. J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, N.J.: Erlbaum, 1981,254-285.

Kieras, D. E. and Polson, P. G. An outline of a theory of the user complexity of devices and systems. Working Paper No. 1, Device Complexity Project, Universities of Arizona and Colorado, May 1982.

Kintsch, W. On modelling comprehension. Invited address at the American Educational Research Association convention. San Francisco, April 10, 1979.

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

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

Miller, J. R., & Kintsch, W. Readability and recall of short prose passages. Text, 1981, 1, 215-232.

Miller, J. R. A knowledge-based mode of prose comprehension: Applications to expository text. Paper presented at the American Educational Research Association meeting, April, 1981.

Kintsch, W., & Miller, J. R. Knowledge-based processes in prose comprehension. Paper presented at the 21st annual meeting of the Psychonomics Society, St. Louis, November, 1980.

Miller, J. R. A knowledge-based model of prose comprehension: Applications to expository text. In B. K. Britton and J. B. Black (Eds.), Understanding expository text. Hillsdale, N. J.: Erlbaum, 1982.

E. Funding Support Status

1. Readability and Comprehension.
   Walter Kintsch, Professor, University of Colorado
   National Institute of Education
   NIE-G-78-0172
   9/1/78 - 8/31/81: $96,627
   9/1/80 - 8/31/81: $46,537

2. Text Comprehension and Memory
   Walter Kintsch, Professor, University of Colorado
   National Institute of Mental Health
   5 Rol MH15872-14-16
   7/1/81 - 6/30/84: $148,693
   7/1/81 - 6/30/82: $37,370

3. Comprehension and Analysis of Information in Text
   Walter Kintsch, Professor, University of Colorado, and
   Lyle E. Bourne, Jr., Professor, University of Colorado
   Office of Naval Research, Personnel and Training Programs
   ONR N00014-78-C-0433
   6/1/78 - 6/30/82: $225,265
   7/18/81 - 6/30/82: $78,125

4. Procedural Net Theories of Human Planning and Problem Solving
   Michael Atwood, Research Psychologist, Science Applications, Incorporated, Denver, Colorado
   Office of Naval Research, Personnel and Training Programs
   ONR N0014-78-C-0165
   1/25/78 - 12/31/80: $230,000
   1/1/80 - 6/30/81: $85,000

5. User Complexity of Devices and Systems
   David Kieras, Associate Professor, University of Arizona
   Peter G. Polson, Professor, University of Colorado
   International Business Machines Corporation
   1/1/82-12/31/82: $104,000

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) recently 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 primary long-term goal of the prose comprehension group is the development of a blackboard-based model of prose comprehension. Correspondingly, we anticipate continued use of the AGE and UNITS packages. These packages allow us to model the knowledge structures possessed by people and the inferential processes that operate upon those structures, and are essential to our work.

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 users 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

The research of the prose comprehension project is clearly tied to continued access to the AGE and UNITS packages, which are simply not available elsewhere. We hope that our continued use of these systems will be offset by the input we have been and will continue to provide to those projects: our relationship has been symbiotic, and we look forward to its continuation.

C. Needs and Plans for Other Computational Resources

We currently use two other computing systems located at the University of Colorado. One is the Department of Psychology's VAX 11/780, which is used primarily to run real-time experiments to be modeled on SUMEX. The second is the University of Colorado's CDC 6400, which is used for various types of statistical analysis.

When the ARPA-sponsored Vax/Interlisp project is completed, we would be most interested in experimenting with becoming a remote AGE/UNITS site. It would seem that this sort of development is the ultimate goal of the package projects, and this type of interaction, once it becomes feasible, would be a logical extension of our association with the SUMEX facility.

D. Recommendations for Future Community and Resource Development

Our primary recommendation for future development within SUMEX involves (a) the continued support of INTERLISP, which is needed for AGE and for other work we have underway on SUMEX and (b) the continued development of the AGE and UNITS projects. In particular, we would like to see an extension of AGE to include a wider variety of control structures so that our psychological models would not be confined to one particular view of knowledge-based processing. The limited physical capacity of SUMEX, both in terms of address space and overloading, is, as before, a major problem. The prose comprehension group can no longer use the publicly released AGE/UNITS system due to its severely limited address space, and has had to build a personal AGE system from a stripped-down version of Interlisp and a selected subset of AGE and UNITS. We heartily endorse the plans underway to obtain more computing capacity for the SUMEX project.

Given our acquisition of a VAX, we particularly support the ongoing and continued development of INTERLISP for the VAX, so that local use of AGE and UNITS would be possible. Since we, as well as other psychologists, need the real-time capability of VAX/VMS to run on-line experiments, we hope that the INTERLISP system to be developed will be compatible with VMS. Note that this need for real-time work coincides with real-world applications of SUMEX programs, in which a VAX might be devoted to both real-time patient monitoring and diagnostic systems such as PUFF or MYCIN.

II.A.2.4 PUFF-VM Project

PUFF-VM: Biomedical Knowledge Engineering in Clinical Medicine

John J. Osborn, M.D.
The Medical Research Institute of San Francisco
Pacific Medical Center

and

Edward H. Shortliffe, M.D., Ph.D.
Department of General Internal Medicine
Stanford University Medical Center
Stanford University

The immediate goal of this project is to develop knowledge-based programs to interpret physiological measurements made in clinical medicine. The interpretations are intended to be used to aid in diagnostic decision making and in therapeutic actions. The programs will operate within medical domains which have well developed measurement technologies and reasonably well understood procedures for interpretation of measured results. The programs are:

(1) PUFF: the interpretation of standard pulmonary function laboratory data which include measured flows, lung volumes, pulmonary diffusion capacity and pulmonary mechanics, and

(2) VM: management of respiratory insufficiency in the intensive care unit.

The second, but equally important, goal of this project is the dissemination of Artificial Intelligence techniques and methodologies to medical communities that are involved in computer aided medical diagnosis and interpretation of patient data.

I. SUMMARY OF RESEARCH PROGRAM

PUFF:

A. Project Rationale

The task of PUFF program is to interpret standard measures of pulmonary function. It is intended that PUFF produce a report for the patient record, explaining the clinical significance of measured test results. PUFF also must provide a diagnosis of the presence and severity of pulmonary disease in terms of measured data, referral diagnosis, and patient characteristics.

B. Medical Relevance and Collaboration

Interpretation of standard pulmonary function tests involves attempting to identify the presence of obstructive airways disease (OAD: indicated by reduced flow rates during forced exhalation), restrictive lung disease (RLD: indicated by reduced lung volumes), and alveolar-capillary diffusion defect (DD: indicated by reduced diffusivity of inhaled CO into the blood). Obstruction and restriction may exist concurrently, and the presence of one mediates the severity of the other. Obstruction of several types can exist. In the laboratory at the Pacific Medical Center (PMC), about 50 parameters are calculated from measurement of lung volumes, flow rates, and diffusion capacity. In addition to these measurements, the physician may also consider patient history and referral diagnosis in interpreting the test results and diagnosing the presence and severity of pulmonary disease.

Currently PUFF contains a set of about 400 physiologically based interpretation "rules". Each rule is of the form "IF <condition> THEN <conclusion>". Each rule relates physiological measurements or states to a conclusion about the physiological significance of the measurement or state.

C. Progress Summary

The results of the PUFF system are reviewed in more detail in the 1978 SUMEX annual report and [3]. A version of the PUFF system is now in routine daily use at Pacific Medical Center. Reports are reviewed by a physician pulmonary physiologist. Over 85 % of the reports are accepted by the physician without change; they are signed and entered into the patient record. Most of the remaining reports are edited on-line to modify a small point in the test interpretation.

The PUFF system was developed on SUMEX in INTERLISP. It was converted to BASIC to run on a PDP11. It has interpreted the data of about 4000 patient cases since it became clinically operational late in 1978. The hospital has paid the Medical Research Institute for each clinical test interpreted by the system since it became fully operational. The program has been extended and maintained by research personnel during this period. PMC is now contracting to have PUFF converted to run on its own stand-alone micro computer (in BASIC). At that time, PUFF will have been fully moved from the research environment to the full responsibility of the end user.

D. Relevant Publications

[1] "A Physiological Rule-Based System for Interpreting Pulmonary Function Test Results", J.C. Kunz, R.J. Fallat, D.H. McClung, B.A. Votteri, J.S. Aikins, H.P. Nii, L.M. Fagan, E.A. Feigenbaum, HPP 78-154, Stanford Heuristic Programming Project, 1978.

[2] "Prototypes: An Approach to Knowledge Representation for Hypothesis Formation", Aikins, J.S., HPP-79-10 (working paper), 1979. Also Int. Joint Conf. on Artif. Intell., Tokyo, Japan, August, 1979.

[3] "A Physiological Rule-Based System for Interpreting Pulmonary Function Test Results", J.C. Kunz, R.J. Fallat, D.H. McClung, B.A. Votteri, J.S. Aikins, H.P. Nii, L.M. Fagan, E.A. Feigenbaum, Proceedings of Computers in Critical Care and Pulmonary Medicine, IEEE Press, 1979.

[4] "The Art of Artificial Intelligence: Themes and Case Studies of Knowledge Engineering", E.A. Feigenbaum, Proceedings of the IJCAI, (1977). (Also Stanford Computer Science Department Memo STAN-CS-77-612).

VM:

A. Project Rationale

The Ventilator Manager program (VM) interprets the clinical significance of time varying quantitative physiological data from patients in the ICU. This data is used to manage patients receiving ventilatory assistance. An extension of a physiological monitoring system, VM (1) provides a summary of the patient's physiological status appropriate for the clinician; (2) recognizes untoward events in the patient/machine system and provides suggestions for corrective action; (3) suggests adjustments to ventilatory therapy based on a long-term assessment of the patient status and therapeutic goals; (4) detects possible measurement errors; and, (5) maintains a set of patient-specific expectations and goals for future evaluation. The program produces interpretations of the physiological measurements over time, using a model of the therapeutic procedures in the ICU and clinical knowledge about the diagnostic implications of the data. These therapeutic guidelines are represented by a knowledge base of rules created by clinicians with extensive ICU experience.

B. Medical Relevance and Collaboration

To assist in the interpretation process, VM must be able to recognize unusual or unexpected clinical events (including machine malfunction) in a manner specifically tailored to the patient in question. The interpretation task is viewed as an ongoing process in the ICU, so that the physiological measurements must be continually reevaluated producing a current clinical picture.

This picture can then be compared with previous summary of patient status to recognize changes in patient condition upon which therapy selection and modifications can be made. The program must also determine when the measurements are most likely to be sensitive to error or when external measurements would be of diagnostic significance.

VM offers a new approach toward more accurate recognition of alarm conditions by utilizing the history and situation of the patient in the analysis. This is in contrast to the use of static limits applied to measurements generated to fit the "typical patient" under normal conditions. Our program uses a model of interpretation process, including the types and levels of conclusions drawn manually from the measurements to provide a summary of patient condition and trends. The program generated conclusions are stated at levels more abstract than the raw data; for example, the presence of hemodynamic stability/instability rather than in terms of heart rate and mean arterial pressure. When the data is not reliable enough to make these conclusions, additional tests may be suggested. The recognition of important conclusion for which external verification is sought, will also elicit the suggestion for confirming tests from the program.

C. Progress Summary

VM is provided with the values of 30 physiological measurements on a 2- or 10-minute bases by an automatic monitoring system. The output is in the form of suggestions to clinicians and periodic summaries . During the past year, the VM system was run extensively to validate the contents of its knowledge base. The knowledge representation and system output are discussed more extensively in the 1981 SUMEX report and in (Fagan 80).

During the past year, VM was tested on several dozen patients. Rule modifications were made to improve system accuracy. Limits of expected values were refined, and rules relating ventilatory management to hemodynamic function were extended.

During the past year, basic research was started on knowledge representation for a physiological model. This research project integrates Artificial Intelligence (AI) and simple mathematics to analyze a physiological model. In the domain of renal physiology, a computer program is being implemented based on these techniques. It analyzes physiological behavior, diagnoses abnormality, and explains the rationale for its analyses. The program fits data to the model, identifies whether the data are abnormal, and identifies the possible causes and effects of any abnormalities. The physiological model is based on knowledge about anatomy, the behavior of the physiological system, and the mechanism of action of the system.

In this project, mathematics is used in an AI system to compute quantitative values based on available data. Concurrently, AI enhances the simple mathematical models. AI is used to represent diverse knowledge about the problem, to match a solution technique with problems needing solution, and to interpret results. Results may be quantitative and qualitative. Results are explained in terms of their contribution to diagnostic and therapeutic decisions.

During this year, work was also initiated to convert alarm rules from VM into computer programs which routinely analyze patient monitoring data. VM rules have been elaborated and recoded to detect instrument error in respiratory and hemodynamic monitoring data. Additional VM rules have been recoded to detect simple physiological abnormalities. One of the original VM goals was to provide prompt feedback to clinicians about the presence of error conditions. Alarm checking rules have been incorporated into the operating data interpretation programs. Integrating these rules with the production monitoring system software allows this feedback to be prompt and associated naturally with the other monitoring system output.

D. Relevant Publications

Fagan, L.M., Kunz, J.C., Feigenbaum, E.A. and Osborn, J.J.: A symbolic processing approach to measurement interpretation in the intensive care unit. Proc. Third Annual Symposium Computer Applications in Medical Care, Silver Spring, Maryland, October, 1979, pp. 30-33.

Fagan, L.M., Shortliffe, E.H. and Buchanan, B.G.: Computer-based medical decision making: From MYCIN to VM. Automedica 3(2), 1980.

Fagan, L.M.: VM: Representing Time-Dependent Relations in a Medical Setting, Ph.D. dissertation, Stanford University, 1980.

Osborn, J.J., Fagan, L.M., Fallat, R.J., et al: Managing the data from respiratory measurements. Med. Instrumentation, November-December, 1979. (Winner of the 'Best Article of the Year' Award for AAMI - 1979.)

Sierra, D.: Development of a smart respiratory alarm system using artificial intelligence (abstract), presented at AAMI meeting, May 1982.

Mitchell, R.R.: The Need for closed loop therapy. Accepted for publication, Critical Care Medicine, 1982.

Mitchell, R.R., Wilson, R.M., Holzapfel, L., Benis, A.M., et al: The Oxygen washing method for monitoring Functional Residual Capacity. Accepted for publication, Critical Care Medicine, 1982.

E. Funding Support

PUFF-VM was supported by NIH grant GM24669 from 1978 - 1981. A renewal application was approved and not funded.

Research is now supported by a $50,000 grant from the Johnson and Johnson foundation, John J. Osborn, PI. This grant includes some indirect costs. The grant is for one year.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Collaborations and Medical Use of Programs via SUMEX

The PUFF-VM project requires very close collaboration between investigators at two institutions separated by fifty miles. This kind of collaboration, in which program development and testing proceeds concurrently on the same application system, requires a computer network facility for sharing of code, data and ideas. SUMEX has been used at PMC for running programs developed concurrently by Stanford and PMC staff, and data has been taken from the PMC computer system and transferred to SUMEX on magnetic tape for program development and testing. The SUMEX staff has developed a cooperating set of computer programs to allow the PMC computer and the SUMEX/2020 systems to actively exchange files and program data and output.

We also use the SUMEX system for purposes other than program development. Joint PMC-Stanford reports on PUFF and VM are being prepared entirely through the the communications and word processing capabilities of SUMEX. Investigators from the two institutions have collaborated in writing reports together; the separate contributions are prepared on SUMEX, edited and merged with an exchange of messages but without ever requiring actual meetings. We have also used the system for trading bibliographic information with other AIM users. We have also experimentally run the Internist program using SUMEX.

B. Sharing and Interactions with Other SUMEX-AIM Projects

We have participated in the AIM workshop and had very fruitful interaction with a number of other SUMEX users, directly influencing our perception of important problems and potentially appropriate solutions. Personal contacts at other conferences, at Stanford AI weekly meetings, and at PMC with visiting members of the AIM community, have also been very helpful in keeping abreast of the current thinking of other members of the AI community and with members of the medical community interested in computer based physiological analysis and diagnosis. We believe that the use of a common machine and the existence of the AIM conference encourages increased recognition and better communication with other AIM workers. Within AIM we most closely collaborate with the MYCIN, MOLGEN and DENDRAL projects, who share common space, common techniques, and common attitudes.

C. Critique of Resource Management

The SUMEX community continues to be an extremely supportive environment in which to do research on uses of artificial intelligence in clinical medicine. The community has two equally vital resources -- the people with knowledge and interest in AI and the facility on which AI system development can proceed. They are equally excellent as resources, helping hands when faced with problems, and friendly support for continued productive research. The availability of INTERLISP; of a facility on which routine data processing functions (eg. manipulating magnetic tapes and making long listings) can take place; and of message-sending among remote users are all vital functions for our project. SUMEX provides them in an environment which is friendly and reliable. Management of the SUMEX facility is consistent and excellent.

III. RESEARCH PLANS

A. Long Range Goals and Plans

The main emphasis of this project changed in 1979 from the development of the PUFF system to the development, extension and evaluation of the VM system. That change was consistent with the transition of PUFF from a research project to an operating program in a clinical setting. In addition, we want to pursue some promising research opportunities in the richer VM setting. Some long term interests, such as consensus building between experts, may be examined using both application areas.

The long range clinical research goal of the VM project is to develop and evaluate an interpretation system that will improve patient care in the ICU. The basic research goal continues to be to develop new Artificial Intelligence techniques which contribute to progress toward that clinical research goal. Within our relatively limited current funding opportunities, we plan to continue research toward these goals. We plan to exploit the knowledge in the VM rules in the current generation of monitoring system software and to develop improved AI techniques for doing diagnosis using models of physiology.

The VM alarm detection rules will be refined and incorporated into the operating monitoring system analysis programs. Some of these rules could not be fully exploited with VM itself because, running at SUMEX, VM typically received data only every 10 minutes. Incorporated into the analysis routines, the alarm detection rules can immediately obtain additional confirming data as necessary. This proximity to the data will allow both improved rules and more effective clinical use of the available rules.

Extending VM rules into the operating data analysis programs is an activity which follows in the tradition of PUFF. The original problem was analyzed and formalized in the computationally rich SUMEX environment. Once the knowledge of the problem is relatively well understood then it can be represented in the comparatively restrictive environments of BASIC on a PDP11 (in the case of PUFF), or the C language on a microcomputer (in the case of the VM alarm rules). This mode of transfer of knowledge seems to be one important way for knowledge developed in AI systems to be brought into wide use.

The near-term goals for the physiological modeling activity are to continue to develop a theory of physiological modeling within an AI system and to continue development of a computer program which demonstrates the power of that theory. The computer program is designed to demonstrate the following processing capabilities: identify the causes and effects of abnormal values of routine physiological measurements; define an abnormality (e.g. oliguria) identifying its possible causes, effects, and associated therapeutic goals; identify the normal physiological consequences of routine physiological behavior, such a drinking a fluid; identify the physiological consequences of abnormal physiological or anatomical situations.