

## III.B PRELIMINARY SUMEX-AIM CPU AUGMENTATION PLAN

## PRELIMINARY PLAN FOR AUGMENTING SUMEX-AIM TO KL-10

We have indicated a budgetary provision of \$132,051 as funds carried forward from year 02 for system augmentation. As discussed in the report (See "SYSTEM DEVELOPMENT - PDP-10 HARDWARE", page 4), we have been observing system performance over the past few months as the user load has increased and find that the response time loading during prime shift approaches saturation. Based on these data and the discussions in Appendix C, we feel that the next most significant bottleneck to system throughput for the AIM community will be in the central processor. We are very concerned about being able to accommodate adequately both the continuing development of the AI performance programs and the exposure of working physicians in a meaningful way. Of course, the NON-prime time loading does not yet approach saturation: we are working on both technical and managerial incentives to encourage use of off-hours (a surprising number of INTERLISP users already are working in the evening and night to get a less loaded machine). However, the very nature of interactive computing in consultative programs is that human beings are involved and because of other commitments for professional people (e.g., physicians), their main load in using the machine will be during prime time. We must then take some action to increase system throughput to assure satisfactory user response time under the increasing prime time load and our judgement is that the most effective augmentation would be in CPU capacity.

Over the past year a potential solution to this problem has emerged with DEC's announcement of the KL-10 processor, scheduled to begin delivery this summer of 1975. The KL-10 is a microprogrammed machine using a high speed cache memory and inherently faster logic to achieve a throughput estimated by DEC to be 2-3 times that of the KI-10. The relationship of DEC's estimates to the performance of a KL-TENEX with many INTERLISP jobs is not easily evaluated, particularly in view of needed additional detailed information described below. However, for a relatively small cost increment (15%) over the capital investment in the facility, an increased throughput of around 2 times can be expected. In a recent example of where the IMSSS system was up-graded from a KA-10 to a KI-10 (1.8 times as fast), the load average for a comparable job mix dropped from 15-20 down to 5-10. In view of a number of uncertainties about this solution at this time which must be resolved, we propose that unobligated year 02 funds be carried forward and held in reserve by BRB until the necessary technical and administrative decisions can be made on a more detailed proposal to be submitted some months hence. At that time, we also anticipate an enlargement of our user base which would simultaneously increase a) the current loading and b) the justification for remedial investment.

From a technical viewpoint, KL-10's have not been delivered yet and available documentation is quite superficial. We do not want serial number 1 as we are heavily committed to providing RELIABLE

community computing resources. The lack of precise documentation makes it difficult to more than cursorily assess the problems in transferring TENEX to the KL-10. At this time it appears that the KL will look initially (i.e., with the initial microcode) like a KI-10 in terms of page faulting and other functions. This would make mounting TENEX quite straightforward. Other questions arise, however, about the management of system loading and diagnostics. DEC's approach with regard to TOPS-10 is to interface the system utility PDP-11/40 (integral to the KL-10) with the new RP-04 disk hardware for this purpose. One of these drives with controller would cost over \$80,000 alone. DEC is still considering our request for information about supporting other disk systems (as the PDP-11 can easily do) and supplying necessary information about the PDP-11 programming to allow us to modify it. Without such necessary technical information we are delayed in doing the detailed planning. Based upon current list prices from DEC, the cost of upgrading to a KL-10 would be about \$150,000. This price is somewhat above the the amount which apparently can be carried forward (additional funds may be available if they do not have to be spent on the remaining ARPANET interface work) but in light of our previous success in working with DEC and achieving a discount, this price may be brought into range of available funds.

## III.C RESOURCE FUNDING

The SUMEX-AIM resource is essentially wholly funded by the Biotechnology Resources Branch [\*]. The various collaborator projects which use SUMEX are independently funded with respect to their manpower and operating expenses. They obtain from SUMEX, without charge, access to the computing and, in most cases, communications facilities in exchange for their participation in the scientific and community building goals of SUMEX.

[\*] Except for the participation by Stanford University in accordance with general cost-sharing, and for assistance to SUMEX by other projects with overlapping aims and interests.

## IV RESOURCE PROJECT DESCRIPTIONS

The following are inputs from the various user projects currently in the SUMEX-AIM community. These project descriptions and comments are the result of a solicitation for contributions sent to each of the project Principal Investigators requesting the following information:

"Please submit for the SUMEX-AIM annual report:

- 1) Updated abstract of project goals and activities. We plan to assemble these on-line as well for community information.
- 2) A summary of project accomplishments over the past year achieved by means of SUMEX. Please include references to any publications which have resulted.
- 3) Comments and assessments of your experiences in interacting with the SUMEX resource. These should include any successes or problems and include technical aspects as well as community or administrative aspects of your collaboration.

It is also important that you indicate to us how your use or non-use of the system corresponded to your expectations and agreements with us at the time you were enrolled as a SUMEX user. This is the time to report extensions of the scope of your project, as well as pilot trials that you may have initiated. We encourage such experimentation without great formality, but we expect that departures that involve significant computer usage will be reported to us at this time for re-review."

The text which follows on the various projects is primarily the responsibility of the indicated project leaders.

## IV.A FORMALLY APPROVED PROJECTS

## IV.A.1 STANFORD USERS

## IV.A.1.a DENDRAL PROJECT

Principal Investigators: Profs. C. Djerassi (Chemistry),  
J. Lederberg (Genetics), and E. Feigenbaum (Comp. Sci.)

(Grant NIH RR-00612-05, 3 years, \$323,933 this year)

## OVERVIEW

The Heuristic DENDRAL project is an application of artificial intelligence to biomedical molecular structure determination problems. Under NIH funding the project has moved significantly closer to making the computer programs and structure elucidation techniques available to a broad community of scientists. This brief report is organized in three parts according to the three major aims of the project: (PART 1) Enhancing the power of the mass spectrometry resource, (PART 2) Developing performance and theory formation programs, and (PART 3) Applying the computer programs and instrumentation to biomedically relevant structure elucidation problems.

The highlight of the period since May 1, 1974, was the project's move to the interactive computing environment of the NIH-funded SUMEX-AIM facility from the batch computing environment of the Stanford Computation Center. Because of this, many scientists outside this university have been able to use the DENDRAL computer programs for their own research. Also, the programs themselves grew in power and scope, and we opened new vistas for collaboration with other research groups. We have been able to make the programs more conversational and thus more helpful to the chemists and biochemists for whom they were developed. Outside users in other research groups also have in SUMEX an easy mechanism for trying out the DENDRAL programs on their own structure elucidation problems. Finally, we have a mechanism for looking at subroutines developed by other research groups in the context of our own programs -- and have incorporated subroutines written, for example, by T. Wipke and by R. Feldmann, into our procedures. The programs and their development are discussed in Part 2, below.

The DENDRAL project, one of the major users of the SUMEX-AIM computer facility, has been forming its own community of remote users. This "exodendral" community has already provided valuable contributions to program development and both the community and contributions are expected to grow at an increased rate. As an example, for the last month for which figures are available (March 1975), the number of CPU hours used by exodendral persons amounted to at least ten percent of the CPU hours used by the entire DENDRAL project which, of course, still reflects heavy CPU-intensive

development efforts. In the last month alone, one new exodendral account representing at least three users has been added to the system, and another four exodendral users have been invited to begin their usage via various "guest" accounts.

Our programs are receiving heavy use from local users and outside users who are investigating mass spectrometry problems for a variety of different compound classes. In addition, new program developments have extended the scope of biomedical structure elucidation problems for which we can provide some computer assistance. Local users include members of Professor Djerassi's group, other chemistry department persons and research groups at the Stanford Medical School. We have recently begun the process of building a community of outside users who can access our programs at SUMEX via TYMNET or ARPANET. Several research groups have expressed considerable interest; we have demonstrated and explained the programs to several groups and we are currently arranging more demonstrations and assisting other people in learning to use SUMEX and the programs from their own laboratories. These applications are discussed in detail in Part 3, below.

#### PART 1: ENHANCING THE POWER OF THE M.S. RESOURCE

Our grant proposal requested funds for significant upgrading of our capabilities in mass spectrometry. The goals of this upgrading were to provide routine high resolution mass spectrometry (HRMS), combined gas chromatography/low resolution mass spectrometry (GC/LRMS) and to develop a combined gas chromatography/high resolution mass spectrometry (GC/HRMS) facility. In addition, this would provide the capability for new experiments in the detection and utilization of data on metastable ions. These capabilities would then be available as required for application to our wider goal, solution of biomedical structure elucidation problems of community of researchers.

The upgrading included several items of hardware and software development, as follows: 1) Acquire stand-alone computer support for the mass spectrometer because existing facilities were inadequate and very expensive; 2) convert existing software, written in the PL/ACME language into FORTRAN so that it would run on the new system; 3) develop new software as required for the demanding task of GC/HRMS; 4) provide hardware and software for semi-automatic acquisition of data on metastable ions. The initial development phase of this upgrading included performance tests to determine the capabilities and limitations of the GC/HRMS system to define the scope of problems to which it can be applied.

#### PART 2: DEVELOPING PERFORMANCE AND THEORY FORMATION PROGRAMS TO ASSIST IN BIOMEDICAL STRUCTURE ELUCIDATION PROBLEMS

The Heuristic DENDRAL computer programs assist with structure elucidation problems by helping interpret mass spectra and helping generate structures that are consistent with the interpretations. The Meta-DENDRAL programs assist with rule formation problems in cases where the rules of mass spectrometry are not known.

All programs have been transferred to the SUMEX machine and most have been considerably improved from their previous versions. The CONGEN and PLANNER programs, in particular, have been improved substantially because these two were thought to offer the most to scientists with structure elucidation problems. Two new programs were developed in this period: CLEANUP and MOLION. The CLEANUP program helps separate the mass spectra of individual components from a GC/MS analysis, and eliminates the background due to GC column "bleed". The MOLION program determines the mass and empirical formula of the whole molecule from its mass spectrum, without prior knowledge of any of the features of the molecule. Both of these new programs solve major problems that had to be determined manually previously before a scientist used the DENDRAL programs.

### PART 3: APPLICATIONS TO BIOMEDICAL STRUCTURE ELUCIDATION PROBLEMS

One of our major aims is to apply the instrumentation and computer programs described above to the study of molecular structure problems in a variety of biomedical applications areas. In order to do that we have made it quite clear that our facilities would be made available to wider community of collaborators/users as our resources permitted. Both categories of application, i.e., within our own group, and with an outside group, are described below.

We have taken several steps toward encouraging a broad community of potential users to call on our facilities. For example, we sent a memorandum to local persons who had indicated their potential need for our facilities as described in our proposal. A questionnaire sent to members of the American Society for Mass Spectrometry, Committee III on Computer Applications, resulted in about 55 persons indicating a desire to know more about access to our programs. A subsequent description of the DENDRAL programs was sent to these persons and to several other persons whom we know from personal contacts might be interested. Because of the nature of their investigations, many of these people receive NIH support.

The availability of SUMEX as a mechanism for resource sharing has made it possible for us to extend access to our programs to a number of people. Without SUMEX, this access would be impossible, and most of our programs (especially LISP routines which are not easily exportable) could be used only by ourselves.

#### Applications by Professor Djerassi's Research Group

Prof. Djerassi's research group at Stanford makes extensive use of the computer programs and the mass spectrometry facility. These applications are detailed in the DENDRAL annual report to the NIH.

#### Applications by Other Members of the Stanford Chemistry Dept.

- 1) Prof. Mosher: We have used CONGEN to suggest structural possibilities for a naturally occurring analog of the fish poison tetrodotoxin. This structure is still under investigation.
- 2) Prof. Hahn (on sabbatical leave at Stanford from Syracuse University): We have used CONGEN to explore possible structures for unknown products of a photochemical reaction. These results have led him to begin a new set of experiments (specifically, CMR) to greatly restrict the possibilities.
- 3) Prof. Johnson: In his wide-ranging syntheses of steroid hormones and other steroids of biological interest, he has studied reactions involving stereo-specific cyclization. We are investigating use of CONGEN for structural analysis under constraints imposed by synthetic cyclization experiments. For example, a previously investigated compound was found to have two structural possibilities. The new possibility could not be differentiated from the assigned structure based on available data.
- 4) Prof. Collman: We have utilized our mass spectrometry facilities to analyze samples in support of his work on oxygen binding to porphyrins (hemoglobin models).
- 5) Prof. Van Tamelen: We have provided mass spectrometry support (HRMS) to assist in the characterization of several compounds related to his work on terpenoid cyclizations.

#### Applications by Other Stanford University Scientists

- 1) Genetics Research Center (GRC) Stanford Hospital: One of our strongest collaborations because of their requirements for additional automation in data reduction and analysis. Their screening program for metabolites characteristic of diseases of genetic origin uses GC/LRMS as the primary source of data. The CLEANUP and MOLION programs were written at least in part to assist the GRC in more systematic approaches to their data. We are currently using CONGEN to assist in determination of structures of unknowns for which mass spectrometric and chemical data are available. Our GC/HRMS facilities will also be utilized for problems which require determination of empirical formulas for ions in spectra of unknown compounds.
- 2) Stanford Pharmacy: We have had several requests for assistance from the Pharmacy of Stanford Hospital (Director: Dr. Hiram Serra). These have variously involved analyzing the stability and purity of pharmaceutical preparations, in particular:
  - a. the impurity of stock preparations
  - b. the stability of nitroglycerine tablets to heat;
  - c. the stability over several months of methyl-dopa, prednisone and banthine when these compounds were formulated into syrups.

- 3) Drug Assay Laboratory Department of Pharmacology, Stanford University: Research personnel from this laboratory (Director: Summer M. Kalman) have requested mass spectra on various derivatives of digoxin using both high and low resolution data.
- 4) Department of Psychiatry, Stanford University: The research group headed by Dr. J. Barchas has used low resolution mass spectral data for the purpose of structure elucidation of a basic compound of interest to their research program.
- 5) Department of Anesthesia, Stanford University: The DENDRAL group was asked by Dr. J. Trudell to help him in the identification of a urinary metabolite isolated after the administration of an anesthetic. This work involved high resolution mass spectrometry of fractions isolated by Dr. Trudell.
- 6) Department of Psychiatry, Palo Alto Veterans Hospital: In this work we analyzed samples by GC/MS given to us by Dr. S. Kanter who works with Dr. Hollister. They were interested in detecting cannabinal, delta-9-tetrahydrocannabinol and an unknown (molecular weight 312) from urine extracts of subjects who had smoked marijuana. This involved running standards of cannabinal and its delta-9-tetrahydro analog through the GC/MS. We were unable to identify these compounds by mass spectrometry as being present in urine. In a subsequent meeting we learned that their concentration was less than 20 nanogram (per GC/MS injection) which is below the limits of sample flow for the recording of reproducible mass spectra. Dr. Kanter is working on the problem of isolating sufficient material for GC/MS and we expect to continue this project in year II of the current grant.
- 7) Prof. McCarty - Civil Engineering: Prof. McCarty is involved in a project to monitor water quality of effluents from tertiary sewage plants. This project includes significant efforts at characterization of the organic content of the water in various phases of its treatment to determine the efficiency of removal of various materials and to identify unknown organic compounds. We have agreed to provide instrumental and computer program support where necessary to assist him in characterization of these samples.

#### Applications by Non-Stanford Scientists

As an additional component of the resource sharing aspects of research, we have, as resources allow, extended the use of our facilities to a group of users remote from the local Stanford community. We have divided these users into two categories, those for whom we have provided mass spectrometry support and those who represent users of DENDRAL programs and collaborators on program development via the SUMEX resources.

#### A. Users of Mass Spectrometry Facilities

- 1) Professor O. O. Orazi, La Plata, Argentina: During the past year we have supplied Dr. Orazi with three low resolution mass spectra. We will be providing HRMS data for him in year II of our grant.
- 2) Professor T. Nakano, Caracas, Venezuela: Dr. Nakano sent one sample of an unknown alkaloid for high resolution mass spectrometry. We were able to show that his low resolution mass spectrum was 2 amu from the true molecular ion and after recording a low resolution mass spectrum his alkaloid was identified as a known compound.
- 3) Dr. Steen Hammerum, Copenhagen, Denmark: Dr. Hammerum requested our assistance in running ultra high resolution mass measurements on several ions in the mass spectra of compounds he had specifically labelled with  $^{13}\text{C}$ .

#### B. Users/Collaborators of/with DENDRAL Programs on SUMEX

Below, in alphabetical order, we list those persons who have a) expressed interest in use of our programs and have been sent instructions in how to gain access to SUMEX and our programs. In many cases these persons have received more detailed information in the form of demonstrations in person or remotely using the LINK facilities of SUMEX, and b) persons who have acted as collaborators in development of parts of one or more of our programs.

Because we have just begun encouraging a significant community of persons to try our programs, we do not yet have a good idea of which persons will continue as serious users. But we have at least provided the opportunity for persons to gain access to our programs, try them and determine how they might (or might not) fit into their own research problems. The term "exploratory" refers precisely to this category of persons who are now engaged in this kind of evaluation. The program names after each person's activity refer to their current major interest. In some cases, we do not actually know the specific problems which are being explored.

1. Dr. A.L. Burlingame (U.C. Berkeley) - Exploratory - all DENDRAL programs.
2. Prof. E.J. Corey (Harvard) - Exploratory, collaboration on programming strategies, CONGEN.
3. Dr. L. Dunham (Zoecon) - Exploratory - Structure determination - CONGEN.
4. Dr. H.M. Fales (NIH) - Exploratory - all DENDRAL programs.
5. R. Feldmann (NIH) - Collaborative development of programs (structure input and drawing routines).
6. Prof. D.L. Fishel (Kent State) - Considering access to SUMEX - has the program descriptions.

7. Prof. M.J. Goldstein (Cornell) - We have provided CONGEN results to him for a difficult structure problem.
8. Dr. N.A.B. Gray (Cambridge) - Collaborating on strategies for computer-assisted structure elucidation programs. He is working on spectral data interpretation.
9. Dr. P. Gund (Merck, Sharpe & Dohme) - Arranging an on-line demonstration for exploratory purposes - CONGEN.
10. Dr. J. Karliner (Ciba-Geigy) - Using CONGEN on structure elucidation problems.
11. Dr. S. Heller (Environmental Protection Agency) - Collaboration on mass spectral library development.
12. Dr. P. Jurs (Penn. State) - Collaboration on structure analysis and building of chemical structure models.
13. Dr. B. Kowalski (Univ. of Washington) - Has approached us for use of SUMEX in pattern recognition work.
14. Dr. D. Lefkowitz (Univ. of Penn.) - Exploratory - interest in DRAW portion of CONGEN for NCI chemical information system.
15. Dr. S. Markey (NIH) - Exploratory - all DENDRAL programs.
16. Dr. F. McLafferty (Cornell) - Exploratory - all DENDRAL Programs.
17. Dr. R. Milberg (National Center for Tox. Res.) - Exploratory - CONGEN.
18. Dr. D. Poulter (Univ. of Utah) - Exploratory - use of CONGEN in structure determination problems, especially terpenoids.
19. Dr. K. Rinehart (Univ. of Illinois) - Exploratory - all DENDRAL programs.
20. Dr. P. Roller (National Cancer Institute) - Exploratory - all DENDRAL programs.
21. Dr. R. Rosen (FMC Corp.) - Exploratory - all DENDRAL programs.
22. Dr. G. Szonyi (Polaroid Corp.) - Interest in CONGEN, exploratory phase beginning.
23. Dr. W.T. Wipke (Princeton) - Exploratory - CONGEN collaboration on structure model building and methods for stereochemical representation of chemical structure.

For a further discussion of the efforts by the DENDRAL project toward network collaboration and dissemination of the programs, see Appendix F which contains a preprint of a paper to be presented at the American Chemical Society symposium on Computer Networking in Chemistry in August of 1975.

#### DENDRAL PUBLICATIONS

- (1) J. Lederberg, "DENDRAL-64 - A System for Computer Construction, Enumeration and Notation of Organic Molecules as Tree Structures and Cyclic Graphs", (technical reports to NASA, also available from the author and summarized in (12)). (1a) Part I. Notational algorithm for tree structures (1964) CR.57029 (1b) Part II. Topology of cyclic graphs (1965) CR.68898 (1c) Part III. Complete chemical graphs; embedding rings in trees (1969)
- (2) J. Lederberg, "Computation of Molecular Formulas for Mass Spectrometry", Holden-Day, Inc. (1964).
- (3) J. Lederberg, "Topological Mapping of Organic Molecules", Proc. Nat. Acad. Sci., 53:1, January 1965, pp. 134-139.
- (4) J. Lederberg, "Systematics of organic molecules, graph topology and Hamilton circuits. A general outline of the DENDRAL system." NASA CR-48899 (1965)
- (5) J. Lederberg, "Hamilton Circuits of Convex Trivalent Polyhedra (up to 18 vertices), Am. Math. Monthly, May 1967.
- (6) G. L. Sutherland, "DENDRAL - A Computer Program for Generating and Filtering Chemical Structures", Stanford Artificial Intelligence Project Memo No. 49, February 1967.
- (7) J. Lederberg and E. A. Feigenbaum, "Mechanization of Inductive Inference in Organic Chemistry", in B. Kleinmuntz (ed) Formal Representations for Human Judgment, (Wiley, 1968) (also Stanford Artificial Intelligence Project Memo No. 54, August 1967).
- (8) J. Lederberg, "Online computation of molecular formulas from mass number." NASA CR-94977 (1968)
- (9) E. A. Feigenbaum and B. G. Buchanan, "Heuristic DENDRAL: A Program for Generating Explanatory Hypotheses in Organic Chemistry", in Proceedings, Hawaii International Conference on System Sciences, B. K. Kinariwala and F. F. Kuo (eds), University of Hawaii Press, 1968.
- (10) B. G. Buchanan, G. L. Sutherland, and E. A. Feigenbaum, "Heuristic DENDRAL: A Program for Generating Explanatory Hypotheses in Organic Chemistry". In Machine Intelligence 4 (B.

- Meltzer and D. Michie, eds) Edinburgh University Press (1969), (also Stanford Artificial Intelligence Project Memo No. 62, July 1968).
- (11) E. A. Feigenbaum, "Artificial Intelligence: Themes in the Second Decade". In Final Supplement to Proceedings of the IFIP68 International Congress, Edinburgh, August 1968 (also Stanford Artificial Intelligence Project Memo No. 67, August 1968).
- (12) J. Lederberg, "Topology of Molecules", in The Mathematical Sciences - A Collection of Essays, (ed.) Committee on Support of Research in the Mathematical Sciences (COSRIMS), National Academy of Sciences - National Research Council, M.I.T. Press, (1969), pp. 37-51.
- (13) G. Sutherland, "Heuristic DENDRAL: A Family of LISP Programs", to appear in D. Bobrow (ed), LISP Applications (also Stanford Artificial Intelligence Project Memo No. 80, March 1969).
- (14) J. Lederberg, G. L. Sutherland, B. G. Buchanan, E. A. Feigenbaum, A. V. Robertson, A. M. Duffield, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference I. The Number of Possible Organic Compounds: Acyclic Structures Containing C, H, O and N". Journal of the American Chemical Society, 91:11 (May 21, 1969).
- (15) A. M. Duffield, A. V. Robertson, C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, and J. Lederberg, "Application of Artificial Intelligence for Chemical Inference II. Interpretation of Low Resolution Mass Spectra of Ketones". Journal of the American Chemical Society, 91:11 (May 21, 1969).
- (16) B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, "Toward an Understanding of Information Processes of Scientific Inference in the Context of Organic Chemistry", in Machine Intelligence 5, (B. Meltzer and D. Michie, eds) Edinburgh University Press (1970), (also Stanford Artificial Intelligence Project Memo No. 99, September 1969).
- (17) J. Lederberg, G. L. Sutherland, B. G. Buchanan, and E. A. Feigenbaum, "A Heuristic Program for Solving a Scientific Inference Problem: Summary of Motivation and Implementation", Stanford Artificial Intelligence Project Memo No. 104, November 1969.
- (18) C. W. Churchman and B. G. Buchanan, "On the Design of Inductive Systems: Some Philosophical Problems". British Journal for the Philosophy of Science, 20 (1969), pp. 311-323.
- (19) G. Schroll, A. M. Duffield, C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, and J. Lederberg, "Application of Artificial Intelligence for Chemical Inference III. Aliphatic Ethers Diagnosed by Their Low Resolution Mass Spectra and NMR Data". Journal of the American Chemical Society, 91:26 (December 17, 1969).

- (20) A. Buchs, A. M. Duffield, G. Schroll, C. Djerassi, A. B. Delfino, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum, and J. Lederberg, "Applications of Artificial Intelligence For Chemical Inference. IV. Saturated Amines Diagnosed by Their Low Resolution Mass Spectra and Nuclear Magnetic Resonance Spectra", *Journal of the American Chemical Society*, 92, 6831 (1970).
- (21) Y.M. Sheikh, A. Buchs, A.B. Delfino, G. Schroll, A.M. Duffield, C. Djerassi, B.G. Buchanan, G.L. Sutherland, E.A. Feigenbaum and J. Lederberg, "Applications of Artificial Intelligence for Chemical Inference V. An Approach to the Computer Generation of Cyclic Structures. Differentiation Between All the Possible Isomeric Ketones of Composition C<sub>6</sub>H<sub>10</sub>O", *Organic Mass Spectrometry*, 4, 493 (1970).
- (22) A. Buchs, A.B. Delfino, A.M. Duffield, C. Djerassi, B.G. Buchanan, E.A. Feigenbaum and J. Lederberg, "Applications of Artificial Intelligence for Chemical Inference VI. Approach to a General Method of Interpreting Low Resolution Mass Spectra with a Computer", *Chem. Acta Helvetica*, 53, 1394 (1970).
- (23) E.A. Feigenbaum, B.G. Buchanan, and J. Lederberg, "On Generality and Problem Solving: A Case Study Using the DENDRAL Program". In *Machine Intelligence 6* (B. Meltzer and D. Michie, eds.) Edinburgh University Press (1971). (Also Stanford Artificial Intelligence Project Memo No. 131.)
- (24) A. Buchs, A.B. Delfino, C. Djerassi, A.M. Duffield, B.G. Buchanan, E.A. Feigenbaum, J. Lederberg, G. Schroll, and G.L. Sutherland, "The Application of Artificial Intelligence in the Interpretation of Low Resolution Mass Spectra", *Advances in Mass Spectrometry*, 5 (1971), 314.
- (25) B.G. Buchanan and J. Lederberg, "The Heuristic DENDRAL Program for Explaining Empirical Data". In proceedings of the IFIP Congress 71, Ljubljana, Yugoslavia (1971). (Also Stanford Artificial Intelligence Project Memo No. 141.)
- (26) B.G. Buchanan, E.A. Feigenbaum, and J. Lederberg, "A Heuristic Programming Study of Theory Formation in Science." In proceedings of the Second International Joint Conference on Artificial Intelligence, Imperial College, London (September, 1971). (Also Stanford Artificial Intelligence Project Memo No. 145.)
- (27) Buchanan, B. G., Duffield, A.M., Robertson, A.V., "An Application of Artificial Intelligence to the Interpretation of Mass Spectra", *Mass Spectrometry Techniques and Appliances*, Edited by George W. A. Milne, John Wiley & Sons, Inc., 1971, p. 121-77.
- (28) D.H. Smith, B.G. Buchanan, R.S. Engelmores, A.M. Duffield, A. Yeo, E.A. Feigenbaum, J. Lederberg, and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference VIII. An approach to the Computer Interpretation of the High Resolution Mass Spectra of Complex Molecules. Structure Elucidation of Estrogenic Steroids", *Journal of the American Chemical Society*, 94, 5962-5971 (1972).

- (29) B.G. Buchanan, E.A. Feigenbaum, and N.S. Sridharan, "Heuristic Theory Formation: Data Interpretation and Rule Formation". In Machine Intelligence 7, Edinburgh University Press (1972).
- (30) Lederberg, J., "Rapid Calculation of Molecular Formulas from Mass Values". Jnl. of Chemical Education, 49, 613 (1972).
- (31) Brown, H., Masinter L., Hjelmeland, L., "Constructive Graph Labeling Using Double Cosets". Discrete Mathematics, 7 (1974), 1-30. (Also Computer Science Memo 318, 1972).
- (32) B. G. Buchanan, Review of Hubert Dreyfus' "What Computers Can't Do: A Critique of Artificial Reason", Computing Reviews (January, 1973). (Also Stanford Artificial Intelligence Project Memo No. 181)
- (33) D. H. Smith, B. G. Buchanan, R. S. Engelmores, H. Adlerkreutz and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference IX. Analysis of Mixtures Without Prior Separation as Illustrated for Estrogens". Journal of the American Chemical Society 95, 6078 (1973).
- (34) D. H. Smith, B. G. Buchanan, W. C. White, E. A. Feigenbaum, C. Djerassi and J. Lederberg, "Applications of Artificial Intelligence for Chemical Inference X. Intsum. A Data Interpretation Program as Applied to the Collected Mass Spectra of Estrogenic Steroids". Tetrahedron, 29, 3117 (1973).
- (35) B. G. Buchanan and N. S. Sridharan, "Rule Formation on Non-Homogeneous Classes of Objects". In proceedings of the Third International Joint Conference on Artificial Intelligence (Stanford, California, August, 1973). (Also Stanford Artificial Intelligence Project Memo No. 215.)
- (36) D. Michie and B.G. Buchanan, "Current Status of the Heuristic DENDRAL Program for Applying Artificial Intelligence to the Interpretation of Mass Spectra". August, 1973. To appear in Computers for Spectroscopy (ed. R.A.G. Carrington) London: Adam Hilger. Also: University of Edinburgh, School of Artificial Intelligence, Experimental Programming Report No. 32 (1973).
- (37) H. Brown and L. Masinter, "An Algorithm for the Construction of the Graphs of Organic Molecules", Discrete Mathematics (in press). (Also Stanford Computer Science Dept. Memo STAN-CS-73-361, May, 1973)
- (38) D.H. Smith, L.M. Masinter and N.S. Sridharan, "Heuristic DENDRAL: Analysis of Molecular Structure," Proceedings of the NATO/CNNA Advanced Study Institute on Computer Representation and Manipulation of Chemical Information (W. T. Wipke, S. Heller, R. Feldmann and E. Hyde, eds.) John Wiley and Sons, Inc., 1974.
- (39) R. Carhart and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference XI: The Analysis of C13 NMR Data for Structure Elucidation of Acyclic Amines", J. Chem. Soc. (Perkin II), 1753 (1973).

- (40) L. Masinter, N.S. Sridharan, R. Carhart and D.H. Smith, "Applications of Artificial Intelligence for Chemical Inference XII: Exhaustive Generation of Cyclic and Acyclic Isomers". Journal of the American Chemical Society, 96 (1974), 7702. (Also Stanford Artificial Intelligence Project Memo No. 216.)
- (41) L. Masinter, N.S. Sridharan, R. Carhart and D.H. Smith, "Applications of Artificial Intelligence for Chemical Inference XIII: An Algorithm for Labelling Chemical Graphs". Journal of the American Chemical Society, 96 (1974), 7714.
- (42) N.S. Sridharan, Computer Generation of Vertex Graphs, Stanford CS Memo STAN-CS-73-381, July, 1973.
- (43) N.S. Sridharan, et.al., A Heuristic Program to Discover Syntheses for Complex Organic Molecules, Stanford CS Memo STAN-CS-73-370, June, 1973. (Also Stanford Artificial Intelligence Project Memo No. 205.)
- (44) N.S. Sridharan, Search Strategies for the Task of Organic Chemical Synthesis, Stanford CS Memo STAN-CS-73-391, October, 1973. (Also Stanford Artificial Intelligence Project Memo No. 217.)
- (45) D. H. Smith, "Applications of Artificial Intelligence for Chemical Inference XIV: The Number of Structural Isomers of  $C^xN^yO^z$ ,  $x + y + z \leq 6$ . An Investigation of Some Intuitions of Chemists."
- (46) D. H. Smith, "Applications of Artificial Intelligence for Chemical Inference XV." In preparation.
- (47) D. H. Smith and R. E. Carhart, "Applications of Artificial Intelligence for Chemical Inference XVI: On Structural Isomerism of Tricyclodecanes." To be submitted to Journal of the American Chemical Society.
- (48) R. G. Dromey, B. G. Buchanan, D. H. Smith, J. Lederberg and C. Djerassi, "Applications of Artificial Intelligence for Chemical Inference XVII: A General Method for Predicting Molecular Ions in Mass Spectra." To appear in Journal of Organic Chemistry, March, 1975.
- (49) B. G. Buchanan, "Scientific Theory Formation by Computer." To appear in Proceedings of NATO Advanced Study Institute on Computer Oriented Learning Processes, 1974, Bonas, France.
- (50) E. A. Feigenbaum, "Computer Applications: Introductory Remarks," in Proceedings of Federation of American Societies for Experimental Biology, Vo. 33, No. 12 (Dec., 1974) 2331-2332.

## IV.A.1.b MYCIN PROJECT

## Computer Based Consultation in Clinical Therapeutics

Prof. S. Cohen, M.D. (Pharmacology) and  
Dr. B. Buchanan (Computer Science)

(Grant HEW HSO-1544-01, 3 years, \$116,734 this year)

The overall objective of the MYCIN project at Stanford is the development of a computer-based system which is capable of using both clinical data and the judgmental knowledge of experts to improve the effectiveness of medical decision making with regard to clinical therapeutics. The work concentrates initially on the use of antimicrobial agents in the treatment of bacteremias.

The present state of the research includes the following accomplishments:

- 1) An interactive program which utilizes data available from the microbiology and clinical chemistry laboratories, plus the physician's response to computer-generated questions, to provide physicians with consultative advice on antimicrobial therapy. Therapy selection takes into account both the patient's infections and the identities of the organisms causing these infections.
- 2) An interactive explanation capability to permit the program to explain all of its actions and reasoning, including, for example, the deduction of the identities of pathologic agents.
- 3) A capability for computer acquisition of judgmental knowledge about those concepts which the program uses in making deductions; this permits experts in the field of infectious disease therapy to teach the system those therapeutic decisions which they find useful in their clinical practices.

Goals for further development of the system include:

- 1) Expansion of the consultation program to deal with infections other than bacteremias.
- 2) Implementation of the system in the clinical setting at the Stanford Hospital.
- 3) Evaluation of the clinical usefulness of the system and of its impact upon the clinical staff and their prescribing habits.
- 4) Expansion of the rule acquisition system to allow experts to introduce new concepts and decision rules which use these concepts.

- 5) Integration of the rule-acquisition system and explanation capabilities to allow experts to enter new decision rules dynamically during a consultation. This will be useful when the program reaches a conclusion which the expert believes to be erroneous, and the explanation facility indicates that it was reached because the system lacked a vital concept or decision rule.
- 6) The development of meta-rules which will be used for expressing strategies for approaching clinical problems.

The techniques for acquisition, representation, and utilization of knowledge, plus considerations of natural language processing, draw upon current research in artificial intelligence.

#### Past Year's Accomplishments

In the past year we have sought to (i) improve the validity of the system's therapeutic advice and increase the scope of its competence; (ii) redesign the control structure to provide a faster, more direct and more general implementation, with increased attention to human engineering; (iii) develop the program's ability to explain its actions to the user; (iv) develop the capability to acquire new rules through interactions with experts, and (v) evaluate the program's competence and performance in a clinical setting.

#### Validity and Scope

Although MYCIN's original focus was directed only at patients with positive blood cultures, the basic methodology was intended to support a much more general approach to the problem. In the past year the system has gained the ability to deal with infections from which the causative pathogen hasn't been isolated (e.g. pneumonia), or which haven't even been cultured (e.g. brain abscess). In broadening the program from bacteremia, we have also acquired the ability to evaluate the meaningfulness of isolates. In addition, the program has been given a sense of time in order to cope with more easily with the order of events in the patient's history, such as a sequence of cultures. This has made its reasoning more powerful by eliminating, for example, the need to ask explicitly about the order of every pair of events.

An extensive review of the program's approach to drug selection has led to a major revision in the basis for therapy selection. The program now deduces the identity of the organism, the infectious disease diagnosis, and the significance of the organism. These are the primary factors in drug selection, with drug toxicity and ecological factors as secondary considerations. The result is a more appropriate, more sharply focussed drug selection that is able to specify dose, route, and duration.

A review of the treatment of various medical concepts has provided a far more comprehensive collection of rules, which now number almost 200. We have revised the program's approach to some concepts, and provided a much more consistent level of knowledge by filling in many gaps. As part of our efforts to introduce the system to clinicians, it was presented at the meeting of the American Federation of Clinical Research in February, 1975 [1].

### Generality and Usability

A comprehensive review and modification of the control structure was undertaken to improve the program's efficiency and generality. The resulting program is smaller, faster, more direct, and yet more general than the original; response times are now no more than 15 seconds, and typically much less, core size is down by 15%. Also, several new capabilities were added to make the program easier to use. The system is now more tolerant of erroneous or inappropriate responses, and can provide a reworded question, along with a list of acceptable answers. In addition, it can recognize responses which are not sufficiently precise, and rephrase its questions accordingly. (A review of design considerations for the system as a whole is in [2]).

### Explanations

We have developed the explanation facility both to insure that the user can understand the rationale behind the system's recommendations, and to educate him, as well. The system currently can explain both the motivation for questions it asks and the source of the conclusions it draws, as well as answer general questions concerning its store of medical knowledge. All of these capabilities were augmented during the past year, some extensively.

The explanation of motivations for questions underwent a major design revision, which resulted in a much more powerful approach based on the program's knowledge of its own control structure and its ability to examine its rules. The user can now fully explore the system's current line of reasoning.

The language understanding capabilities of the question answering system have also been revised, allowing a broader range of questions to be asked and offering more precise answers. The use of this feature has also been simplified somewhat, so that the user no longer needs to classify all his questions.

### Knowledge Acquisition

A preliminary knowledge acquisition system was completed in the middle of the year. We have demonstrated that a physician can teach the system new rules in a rather stylized subset of English. (An overview of the capabilities is in [3]). Building on the experience gained here, a redesigned system is currently being constructed, which allows the user to examine and modify the program's knowledge and

behavior as a single, unified action. That is, the functions of two separate modules have been combined in a single, redesigned system, that will make use of the fact that the nature of the explanations requested can give a clear hint as to the content of the new rule. It will also advise the user as to the effect of his rule on the original deficiency -- i.e. whether or not it corrects the problem he noticed. Progress has also been made on the issue of assuring the consistency of the knowledge base by examining the nature of contradictions as they appear in our multi-valued logic system.

### Evaluation

Much of the focus for the development in the past year was suggested by the results of a preliminary evaluation done in May, 1974 [4]. This concentrated on the validity and acceptability of the program's advice, by comparing its performance on fifteen cases to that of five infectious disease experts reviewing the same cases. Even at this early stage of development, the results were quite encouraging. None of questions generated by the program were judged irrelevant by a majority of the physicians. In its final diagnosis, the program offered an average of 4.0 possible identities for each organism, while the experts suggested an average of 5.3 possibilities. In almost 50% of the cases, the program's therapeutic regimen was identical to that suggested by the physicians, and 72% of the time it was judged an acceptable alternative. (Agreement among experts was about the same as agreement with the program.)

Another, much more comprehensive study is just beginning, with the primary focus on the validity of the program's medical advice, its impact on the prescribing habits of the user community, and its acceptance by clinicians, as indicated by their use of the system.

### Comments on SUMEX

One of the most important elements in our attempt to develop a truly competent program is the repeated testing of performance on a wide range of cases. The shared resource concept of SUMEX, with the availability of the system to numerous research groups, has become a significant aid in this testing process. We have begun to receive useful feedback from groups at the University of Rochester and the University of Virginia, both of which have been using the program.

The SNDMSG and other system communication facilities have made communication simple, direct, and effective, despite the large distances involved.

We anticipate that, with the availability of a new and redesigned explanation and knowledge acquisition system (planned for mid-summer), we can widen the scope of this interaction. Rather than having users report bugs to us, requesting information, and our answering their questions and fixing the problems, the system itself

should be capable of handling a subset of such interactions. This will help to put the user in direct touch with the knowledge base of the program, so that he can modify or augment it directly. Thus, in addition to gathering the usual sort of feedback from user experience in running the program, we hope to benefit directly from the expertise of infectious disease experts in various centers across the country.

#### References

- [1] E H Shortliffe, F Rhame, et. al., "MYCIN, A Computer Program providing Antimicrobial Therapy Recommendations", Clinical Research, vol 23, p 107A (abstract) 1975.
- [2] E. H. Shortliffe, S. G. Axline, B. G. Buchanan, S. N. Cohen, "Design Considerations for a Program to Provide Consultations in Clinical Therapeutics". Presented at San Diego Biomedical Symposium 1974 (February 6-8, 1974).
- [3] E. H. Shortliffe, R. Davis, S. G. Axline, B. G. Buchanan, C. C. Green and S. N. Cohen, "Computer-Based Consultations in Clinical Therapeutics: Explanation and Rule Acquisition Capabilities of the MYCIN System", to appear in Computers and Biomedical Research, June 1975
- [4] E H Shortliffe, "MYCIN, A Rule Based Computer Program..." STAN-CS-74-465 Computer Science Department, Stanford University 1974
- [5] E. H. Shortliffe, S. G. Axline, B. G. Buchanan, T. C. Merigan, and S. N. Cohen, "An Artificial Intelligence Program to Advise Physicians Regarding Antimicrobial Therapy." Computers and Biomedical Research, 6 (1973), 544-560.
- [6] E. H. Shortliffe and B. G. Buchanan, "A Model of Inexact Reasoning in Medicine, July, 1974. To appear in Mathematical Biosciences.

## IV.A.1.c PROTEIN STRUCTURE MODELING PROJECT

## Protein Crystallography Project

Dr. S. Freer (Chemistry, U. C. San Diego) and  
Prof. E. Feigenbaum and Dr. R. Engelman (Comp. Sci., Stanford)

(Grant NSF DCR74-23461, 2 years, \$88,436 total)

## I. General Objectives

The protein crystallography project involves scientists at two different universities, pooling their respective talents in protein crystallography and computer science, and using the SUMEX-AIM facility as the central repository for programs, data and other information of common interest. The two groups involved are members of the Computer Science Department at Stanford (Prof. Edward Feigenbaum and Dr. Robert Engelman) and members of the Department of Chemistry at the University of California at San Diego (Prof. Joseph Kraut and Dr. Stephan Freer). The general objective of the project is to apply problem solving techniques, which have emerged from artificial intelligence research, to the well known "phase problem" of x-ray crystallography, in order to determine the three-dimensional structures of proteins. The work is intended to be of practical as well as theoretical value to both computer science (particular AI research) and protein crystallography. Viewed as an AI problem, the objectives are:

1. To discover from expert protein crystallographers the knowledge and heuristics which could be used to infer the tertiary structure of proteins from x-ray crystallographic data, and to formalize this expertise in appropriate data structures and heuristic procedures.
2. To discover a program organization and a set of representations which will allow the knowledge and the heuristics to cooperate in making the search efficient, i.e., generating plausible candidates in a reasonable time.
3. To implement the above in a system of computer programs, the competence of which will have a noticeable impact upon the discipline of protein crystallography.

As a research task in protein crystallography, the objective is to develop a computational system which can infer the tertiary structure of a protein molecule in the absence of phase information normally obtained from multiple isomorphous replacement procedures.

## II. Specific Objectives

We are attacking the phase problem in protein crystallography by

developing a system of computer programs that will enable us to make progress in structure determination without use of multiple isomorphous replacements. Protein crystal structures can be divided into three classes according to how much of the structure is known at the outset of the investigation. In order of increasing difficulty these are:

- I. three-dimensional structure already known, but from crystals in a different space group;
- II. three-dimensional structure of a homologous protein is known;
- III. only the amino acid sequence is known or can be approximately inferred from homologous proteins.

We have selected four proteins for study, which fall into the three classes as shown below:

I	II	III
Structure known	Homologous Structure Known	Amino Acid sequence only
Chymotrypsinogen (type D crystals)	Cytochrome f from Spirula  Cytochrome c2 from rhps. capsulata	Two-iron ferredoxin from Anacystis nidulans

For proteins in classes I and II we have used or plan to use Patterson search techniques and the rotation function of Rossmann and Blow. Our objective here is to directly compare the relative efficacy of the Patterson search and rotation function methods. The output from either of these methods is the orientation and location of a trial structure derived from the crystal structure of the identical or homologous protein (say the backbone atoms plus the beta-carbon atoms) in the unit cell of the unknown crystal structure. The correctness of the trial structure can be verified by crystallographic refinement (Freer, Alden, Carter and Kraut, J. Biol. Chem., v. 249, pp. - ). In order to make refinement a more useful tool for this verification we will introduce techniques for incorporating the amorphous solvent into the trial structure, in hopes of speeding the convergence of the refinement process.

The more difficult class III structure determination of the two-iron ferredoxin protein will be attacked by parallel application of Patterson search techniques and Fourier Bootstrap methods. In the Patterson search investigation we will first try to locate the iron atoms with an anomalous dispersion Patterson and then attempt to determine the orientation of the entire iron-sulfur group by Patterson search in which the iron-sulfur portion of the  $[\text{FeS}(\text{SCH}_2)_2\text{C}_6\text{H}_4]_2(2-)$  synthetic analogue described by Mayerle, et al. (Proc. Nat. Acad.

Sci., v. 79, p. 2429 (1973)) is used as the search group. In addition we will determine from known sequence structure correlations what if any helical regions are strongly predicted by the amino acid sequence and will then search the Patterson with Fe-helix vectors to confirm their existence and determine their position and orientation. We will continue this approach to build up as much of the structure as possible.

The Fourier Bootstrap method begins as a trial and error search in which a sphere is used as the trial structure (Kraut, Biochem. Biophys. Acta, v. 30, p. 264 (1958)). Classical Fourier refinement will be used to reveal the shape of the protein and locate its position with more precision. The refinement will then proceed by direct modification of the electron-density map (eliminating regions of negative density, etc.) followed by Fourier inversion of the modified map, a technique used at UCSD from time to time during the last ten years and which has been employed recently with success by Collins and Legg (Amer. Cryst. Assoc., program and abstracts of summer meeting, v. 261 (1974)) to refine the phases and extend the resolution of the rubredoxin crystal structure.

### III. Summary of Project Accomplishments

Our efforts during the past year have been mainly in building the tools which will be the "expert" components in our overall structure elucidation system. The main accomplishments are listed below.

1. Conversion of the Patterson Search program, PSRCH, to SUMEX. This program is now operational.
2. Development of the Superposition program to infer positions of additional atoms when the coordinates of a partial structure are known. This program is now operational.
3. Conversion of the Rossmann rotation function program to SUMEX. This program was obtained from Rossmann's laboratory at Purdue University. Although the conversion to FORTRAN-10 was straightforward, it has taken several months, even with help from one of the program's authors, to learn to obtain good results from it on test problems. It now performs in some cases at a level similar to the Patterson search program. However, attempts to find non-crystallographic symmetries in real crystal protein structures have thus far been unsuccessful. Testing and debugging of the program is still in progress. Unfortunately, the crystallography group at Purdue has no routine access to computing terminals, so that real-time consulting (see below) has not been exploited.
4. IDATA2, a program for computing theoretical structure factors, given the atomic coordinates of a complete molecule, was transferred from UCSD, where it operates on a CDC-3600, to SUMEX and made operational on SUMEX. As originally received, IDATA2 was

a non-standard Fortran program, using many subroutines written expressly for the CDC-3600 either in a CDC dialect of Fortran or in assembly language. Conversion of the program to SUMEX was considerably expedited by the existence of the TYMNET connection to UCSD, and the linking facility in TENEX. Dr. Stephan Freer, one of the authors of IDATA2, spent several hours at a terminal in his office at UCSD, logged in to SUMEX and linked to Dr. Robert Engelmores at Stanford, while Engelmores went through many iterations of editing, compiling, and (eventually) executing the program. Freer's real-time consulting (a further example of which is given in the next section) was as valuable as if he were working in the same office as Engelmores, and compressed the program conversion time from several weeks to several days.

5. Several ancillary utility programs for translating and rotating sets of atom coordinates, determining structural parameters, reading and writing files containing structure factors, Patterson tables, etc. have been implemented on SUMEX.

#### IV. Interaction with the SUMEX resource

This project was conceived from the outset with the expectation of using a common computing facility, available to one or both groups via remote terminals connected to the facility over a high speed network, such as the ARPANET. Preliminary work began before such a facility existed, and it was necessary to make frequent and lengthy visits to UCSD to obtain the necessary expertise and guidance. Moreover, the first computer programs had to run on two very different computers (a CDC 3600 and an IBM 360/67) so that program development could continue at either university. Most of the time interaction between the two research groups consisted of occasional phone calls or letters. There was typically a turn-around time of several days between a suggestion for debugging a program, or a formulation of a series of computer runs, and the corresponding results. By that time the person who had suggested the work originally had forgotten why, and needed to be brought up to speed again.

All that is past history. All program development, and most communications are now effected on the SUMEX computer. The UCSD group has a direct connection to SUMEX via the TYMNET and ARPANET computing networks. Routine daily communications now take place using the system's message facility. Program files are equally accessible from Stanford and UCSD, so that either group can construct, edit or exercise the programs. Large data files can be transmitted to and from either site via the ARPANET, obviating the ubiquitous problems of machine incompatibilities, difficulties of reading tapes written at foreign sites, etc.

But perhaps the most noticeable change in our modus operandi that has occurred with the advent of computer networking has been the closer coupling between the Stanford and UCSD groups in all phases of the research activity. The following transcript (edited for brevity) of a terminal session illustrates how the system is now used to provide "real time" consultation. A member of the Stanford group