The EtherTIP software has undergone further enhancements in the past year. Portions of this work was done in conjunction with the Stanford Computer Science Department. Among those enhancements are the following:

a. It now accepts incoming connections to line printer ports, and for remote system diagnosis.

b. It can simulate the "old" Stanford EtherTIP for users who have not yet made the transition to the new environment.

c. The user interface is more flexible to suit the needs of an increasingly diverse user community.

The EtherTIP software has developed into a very stable system, and one enjoying good use within the SUMEX community.

5. *10 MB/SEC Ethernet Development* -- SUMEX made a major move this past year to begin our transfer to a 10 megabit/sec network. While the current 3 megabit/sec network continues to serve us well, many new workstations and printers are coming on the market with only 10 MB/SEC interfaces, and in addition, since 3 MB/SEC networks were only used a very few selected settings, it is becoming increasingly difficult to find replacement parts when failures do occur.

Therefore, this past year saw several efforts involved in installing and supporting the SUMEX 10MB/SEC Ethernet ;

a. Reworking the entire Ethernet system software to handle *both* 3 and 10 megabit link level standards, i.e., addressing and encapsulation are transparent to the user levels. We similarly made the network link level protocols transparent to the the user level software. In this way one can communicated using PUP protocols on a 10MB/SEC ethernet and the user software does not have to change.

b. Adding address resolution protocols for PUP and IP so that the 3MB/SEC byte addresses can be translated to 10 MB/SEC hardware addresses for the link level. This enables one to communicate using PUP or IP between 3 and 10 megabit hosts.

c. Integrating XNS and IP into the PUP routing mechanism.

d. Solving some rather subtle software/hardware integration problems in order to simulate "ethernet" on the HPP/Welch Road "twisted pair" ethernet.

e. Bringing up the 3 MB/SEC EtherTIP on the 10 MB/SEC network was a proof that the above worked. It was done without any changes to the TIP software itself by simply relinking it with the 10 MB/SEC system software. This required only one additional piece of logic. When a 10 MB/SEC host wants to communicate using PUP which is a 3 MB/SEC protocol, then it must find its PUP address from some host on the 10 MB/SEC network. The gateway maintains a translation table, and listens for such requests, thus translating the 10 MB/SEC hardware address into a "soft PUP address," and replying to the requesting host.

6. *HPP-SUMEX Communication link* -- The Heuristic Programing Project

(HPP) relocated from its campus location to 701 Welch Road, adjacent to the Stanford campus. Since this group is a primary user of the SUMEX computer facility and the principal focus for core AI research, a communication link between the new location and SUMEX machine room was imperative. Several communication schemes for establishing a reliable and relatively fast link were considered, namely ; microwave, laser, infrared, direct ethernet (by trenching and placing a direct ethernet cable), ATT's T1 service and others.

All of the above schemes would have necessitated large budgetary outlays and some would have imposed lengthy time delays (getting permits and the like) due to jurisdictional boundaries. The idea of using bare copper telephone pair already in place looked very attractive especially if reasonable speed and reliability could be achieved. The wire distance between the above mentioned locations is approximately 2000 ft. A design goal was established to try to develop a communication link with Ethernet type speed (3MB/SEC) between these two locations.

Utilizing high driving capacity drivers (differential) and ultra high speed, high sensitivity receivers a transceiver was designed and tested for maximum transmission speed with maximum reliability. The final configuration resulted in a half duplex transmission over a bare copper twisted pair in each direction utilizing Manchester coding at a reliable transmission speed of 1.25MBs/sec. each direction for an aggregate speed of 2.5MBs/sec. This communication link has been in operation for about six months now without any appreciable down time or noticeable error rate or data delays. Many HPP researchers are utilizing this link to communicate with SUMEX and the University Ethernet network. In addition, various Lisp machines and printers located in the HPP facility and connected to a local network there are also able to communicate with the University network.

INTERNET SOFTWARE

One major issue we face at SUMEX-AIM in support of our network environment is the lack of standardization in network protocols among various vendors. Currently, many vendors are adding support to their products for the Internet (IP/TCP) protocols. SUMEX continues to support the IP/TCP protocols on the DEC2060, and we are currently alpha-testing a release of Interlisp-D which also supports IP/TCP protocols. In addition, we sucessfully adapted the IP/TCP software to our VAX systems running UNIX 4.2BSD. This Vax TCP adaptation involved provisions for subnet routing, 3 MB/SEC byte swap problems, encapsulation problems and 10 MB/SEC debugging with our gateways.

I.A.2.6. Progress in Core Research

Over the past year we have continued to support several core research activities aimed at developing information resources, basic AI research, and tools of general interest to the SUMEX-AIM community. SUMEX is providing only partial support for these projects, with complementary funding coming from ARPA, ONR, NLM and NSF contracts and grants to the Stanford Heuristic Programming Project.

Core Research

Core Research at SUMEX-AIM focuses on understanding the roles of knowledge in symbolic problem solving systems, its representation in software and hardware, its use for inference, and its acquisition. We are continuing to develop new tools for system builders and to improve old ones. The research crosses a number of application domains, as reflected in the subprojects discussed earlier, but the main issues that we are addressing in this research are those fundamental to all aspects of AI. We believe this core research is broadening and deepening the groundwork for the design and construction of even more capable and effective biomedical systems.

As mentioned above, although our style of research is largely empirical, the questions we are addressing are fundamental. The three major research issues in AI have, since its beginning, been knowledge representation, control of inference (search), and learning. Within these topics, we will be asking the following kinds of questions. As our work progresses, we hope to leave behind several prototype systems that can be developed by others in the medical community.

1. Knowledge Representation -- How can we represent causal models and structural information? What are the relative benefits of logic-based, rule-based, and frame-based systems? How can we represent temporal relations and events so that reasoning over time is efficient?
2. Knowledge Acquisition -- How can an expert system acquire new knowledge without consuming substantial time from experts? Can we improve the knowledge engineering paradigm enough to make a difference? Can automatic learning programs be designed that will work across many disciplines? Will cooperative man-machine systems be able to open the communication channel between expert and expert system?
3. Knowledge Utilization -- By what inference methods can a variety of sources of knowledge of diverse types be made to contribute jointly and efficiently toward solutions? What is the nature of strategy and control information?

Plans for the Coming Year

Several systems have been developed in recent years to serve as vehicles for knowledge engineering and research on knowledge representation and its use. Knowledge acquisition (including machine learning) and advanced architectures for AI will be the two areas of most new activity in the coming year. Research on these topics obviously must draw on on-going work in representation and control.

In particular, we will focus on

- Inductive learning of MYCIN-like rules from case data in the domain of diagnosing disorders where the chief complaint is jaundice;
- Learning from experience in domains where the means for interpreting new

data are largely contained in the emerging (and thus incomplete and not wholly correct) theory;

- Learning by watching a medical expert diagnose cases presented by NEOMYCIN;
- Investigating complex signal understanding systems for ways to exploit and represent concurrency with a view toward hardware and software architectures that may be capable of several orders of magnitude improvement in performance.

Further information on the core research at SUMEX-AIM and the Heuristic Programming Project can be found in the Projects section starting on page 89.

I.A.2.7. Resource Operations Statistics

The following data give an overview of various aspects of SUMEX-AIM resource usage. There are 5 subsections containing data respectively for:

1. Overall resource loading data (page 31).

2. Relative system loading by community (page 33).

3. Individual project and community usage (page 36).

4. Network usage data (page 44).

5. System reliability data (page 44).

For the most part, the data used for these plots covers the entire span of the SUMEX-AIM project. This includes data from both the TENEX KI10 system and the current DECsystem 2060. At the point where the SUMEX-AIM community switched over to the 2060 (February, 1983), you will notice severe changes in most of the graphs. This is due to many reasons which I will mentioned briefly here ;

1. Even though the Tenex operating system used on the KI10 was a forerunner of the current Tops20 operating system, the Tops20 system is still different from Tenex is many ways. Tops20 uses a radically different job scheduling mechanism, different methods for computing monitor statistics, different I/O routines, etc. In general, it can not be assumed that statistics measured on the Tenex system correlate one to one with similar statistics under Tops20.

2. The KL10 processor on the 2060 is a faster processor than the KI10 processor used previously. Hence, a job running on the KL10 will use less CPU time than the same job running on the KI10. This aspect is further complicated by the fact that the SUMEX KI10 system was a dual processor system.

3. The SUMEX-AIM Community was changing during the time of the transfer to the 2060. The usage of the GENET community on SUMEX had just been phased out. This part of the community accounted for much of the CPU time used by the AIM community. Since the purchase of the 2060 was partially funded by the Heuristic Programming Project (HPP), an additional number of HPP Core Research Projects started using the 2060, increasing the Stanford communities usage of the machine. And finally, the move to the 2060 occurred during a pivotal time in the community when more and more projects were either moving to their own local timesharing machines, or onto specialized Lisp workstations. It also was the time for the closure of many long time SUMEX-AIM projects, like Dendral and Puff/VM.

Any conclusions reached by comparing the data before and after February, 1983 should be done with caution. The data is included in this years annual report mostly for casual comparison. Starting next year, only data from the 2060 will be recorded on the annual report. Readers will be referred to previous annual reports (such as this one) for data from the KI10 Tenex system.

Overall Resource Loading Data

The following plots display several different aspects of system loading over the life of the project. This data includes usage of the Tenex KI10 system and the current DECsystem 2060.

These plots include total CPU time delivered per month, the peak number of jobs logged in, and the peak load average. The monthly "peak" value of a given variable is the average of the daily peak values for that variable during the month. Thus, these "peak" values are representative of ***average*** monthly loading maxima and do not reflect the largest excursions seen on individual days, which are much higher.

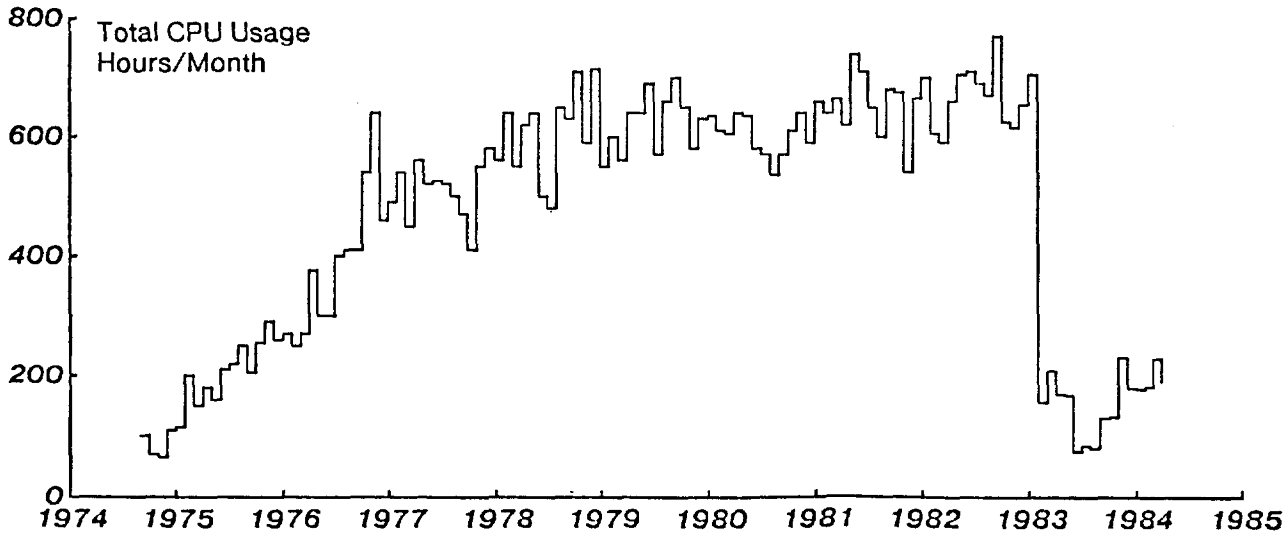

Figure 7: Total CPU Time Consumed by Month

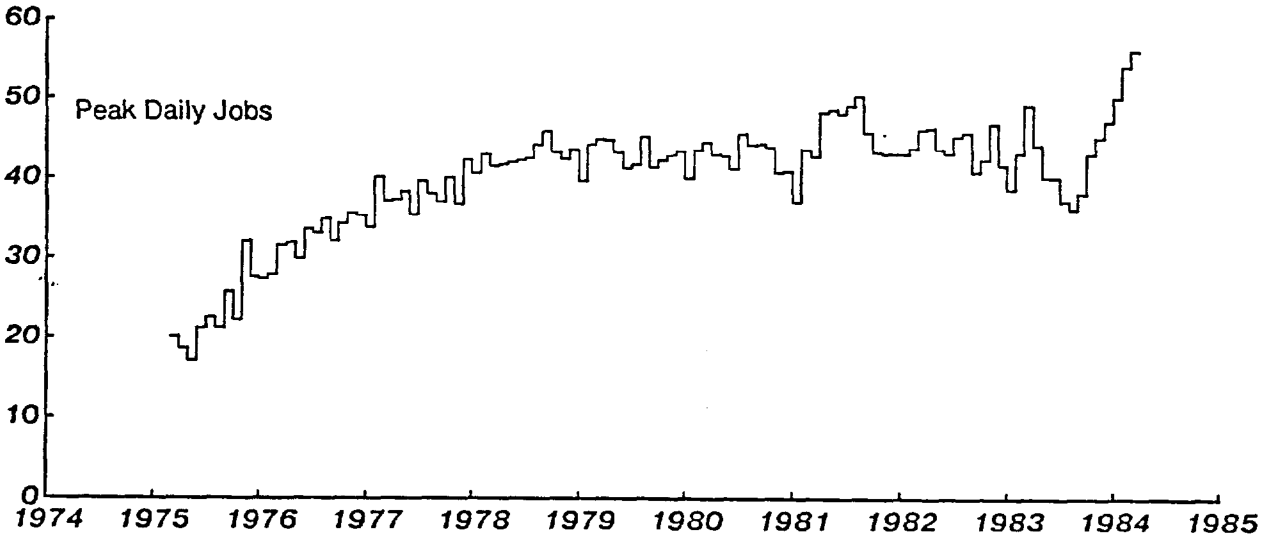

Figure 8: Peak Number of Jobs by Month

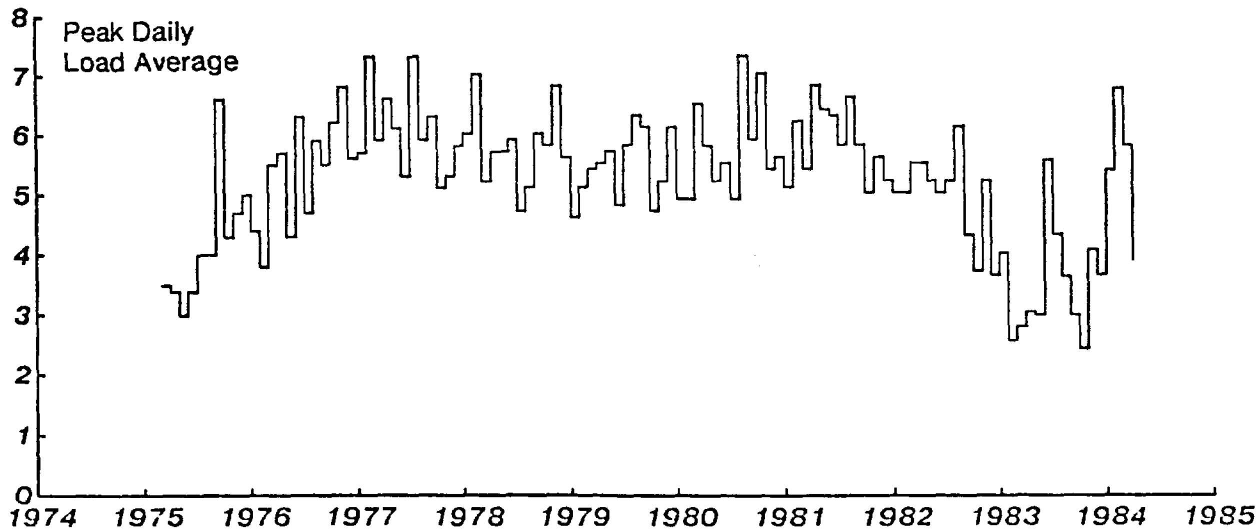

Figure 9: Peak Load Average by Month

Relative System Loading by Community

The SUMEX resource is divided, for administrative purposes, into three major communities: user projects based at the Stanford Medical School (***Stanford Projects***), user projects based outside of Stanford (***National AIM Projects***), and common system development efforts (***System Staff***). As defined in the resource management plan approved by the BRP at the start of the project, the available system CPU capacity and file space resources are divided between these communities as follows:

Stanford	40%
AIM	40%
Staff	20%

The "available" resources to be divided up in this way are those remaining after various monitor and community-wide functions are accounted for. These include such things as job scheduling, overhead, network service, file space for subsystems, documentation, etc.

The monthly usage of CPU resources and terminal connect time for each of these three communities relative to their respective aliquots is shown in the plots in Figure 10 and Figure 11. As mentioned on page 30, these plots include both KI10 and 2060 usage data.

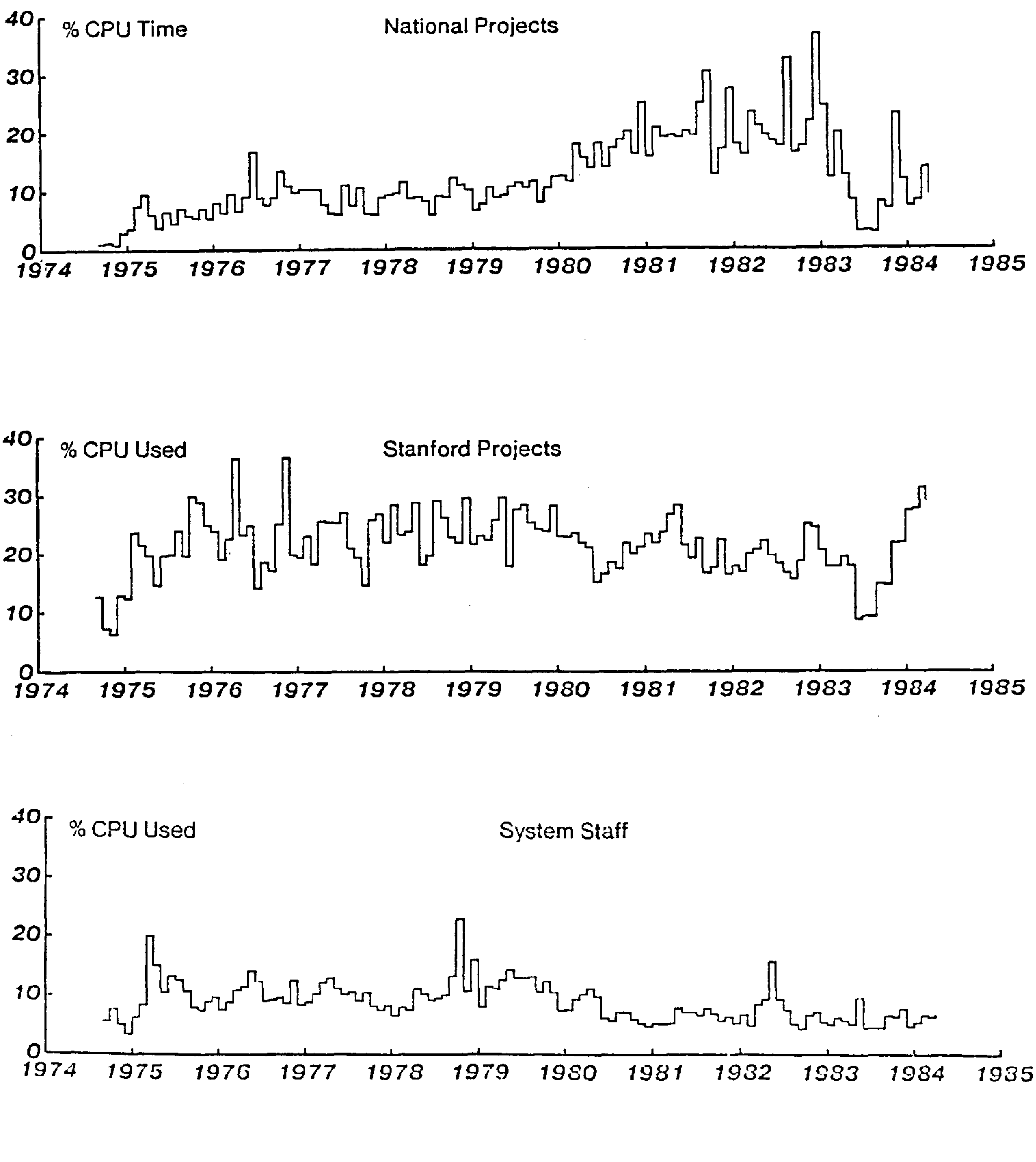

Figure 10: Monthly CPU Usage by Community

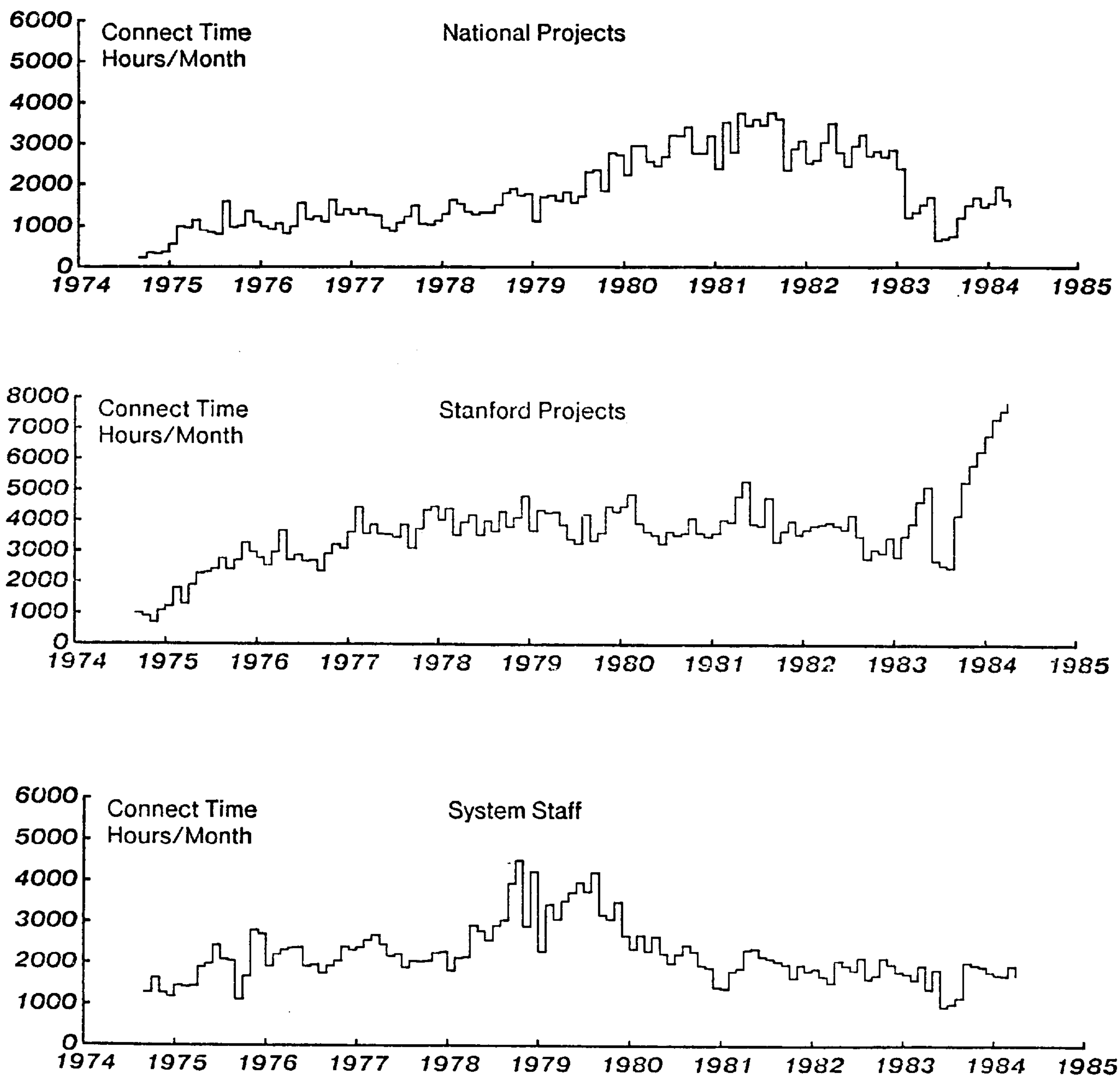

Figure 11: Monthly Terminal Connect Time by Community

Individual Project and Community Usage

The following histogram and table show cumulative resource usage by collaborative project and community during the past grant year. The histogram displays the project distribution of the total CPU time consumed between May 1, 1983 and April 30, 1984, on the SUMEX-AIM DECsystem2060 system.

In the table following, entries include a text summary of the funding sources (outside of SUMEX-supplied computing resources) for currently active projects, total CPU consumption by project (Hours), total terminal connect time by project (Hours), and average file space in use by project (Pages, 1 page = 512 computer words). These data were accumulated for each project for the months between May, 1983 and May, 1984.

Several of the projects admitted to the National AIM community use the Rutgers-AIM resource as their home base. We do not explicitly list these projects in this annual report covering the Stanford SUMEX-AIM resource. We do record information about the Rutgers resource itself, however, and note its separate resource status with the flag "[Rutgers-AIM]".

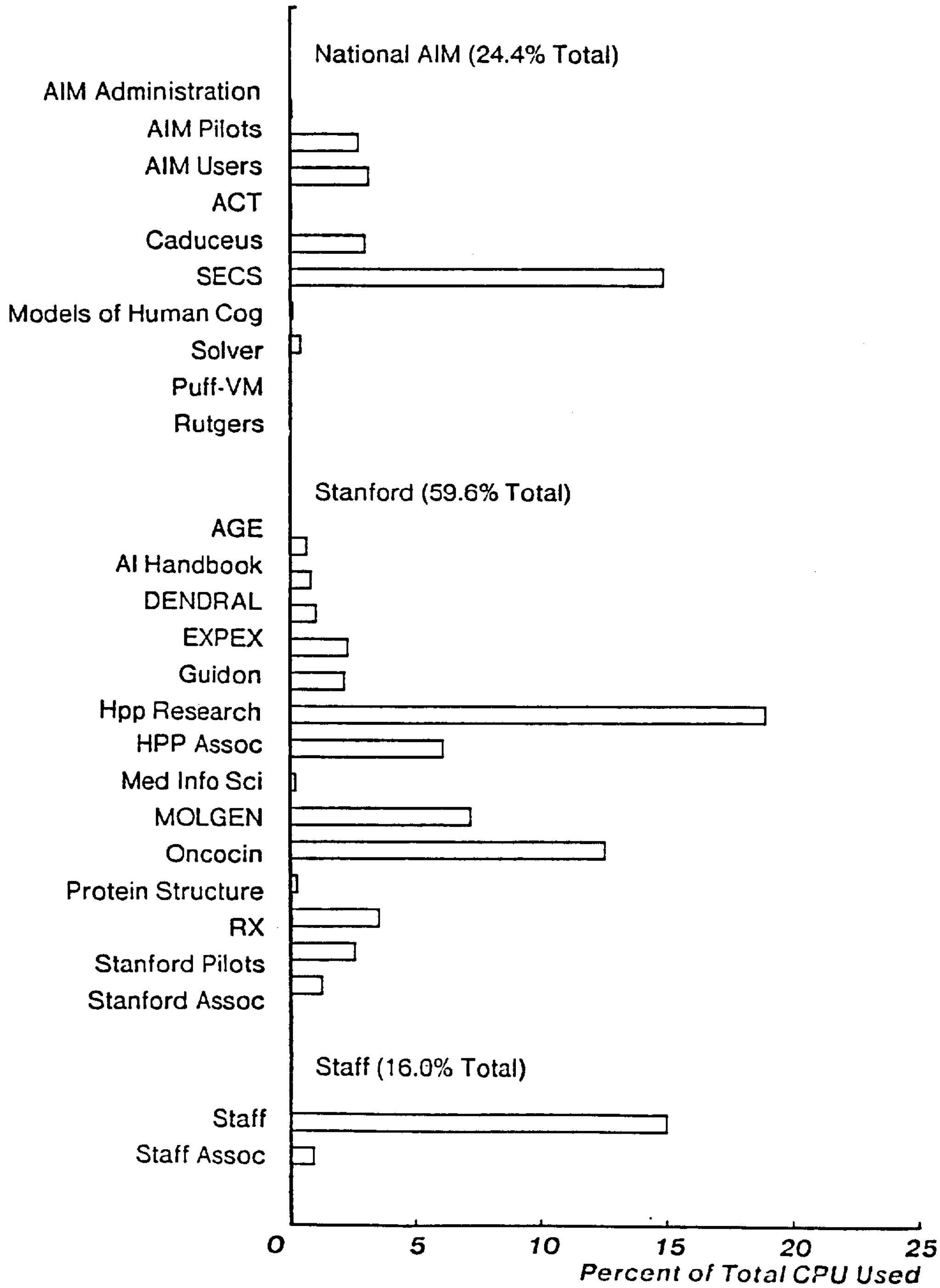

Figure 12: Cumulative CPU Usage Histogram by Project and Community

Resource Use by Individual Project - 5/83 through 4/84

National AIM Community	CPU (Hours)	Connect (Hours)	File Space (**) (Pages)
1) ACT Project "Acquisition of Cognitive Procedures" John R. Anderson, Ph.D. Carnegie-Mellon Univ. NSF IST-80-15357 2/81-2/84 $186,000	0.37	33.88	2866
2) CADUCEUS "Clinical Decision Systems Research Resource" Jack D. Myers, M.D. Harry E. Pople, Jr., Ph.D. University of Pittsburgh NIH RR-01101-07 7/80-6/85 $1,607,717 7/83-6/84 $369,484 NLM LM03710-04 7/80-6/85 $817,884 7/83-6/84 $196,710 NLM New Invest LM03889-02 Gordon E. Banks, M.D. 4/82-3/85 $107,675 4/83-3/84 $35,975 4/84-3/85 $35,975	58.15	895.52	6852
3) CLIPR Project "Hierarchical Models of Human Cognition" Walter Kintsch, Ph.D. Peter G. Polson, Ph.D. University of Colorado NIMH MH-15872-14-16 (Kintsch) 7/81-6/84 $281,085 7/83-6/84 $69,878 NSF (Kintsch) 8/83-7/86 $200,000 IBM (Polson) David Kieras University of Arizona 1/82-12/84 $364,000 1/84-12/84 $145,000	1.38	209.34	750

4) PUFF-VM Project "Biomedical Knowledge Engineering in Clinical Medicine" John J. Osborn, M.D. Med. Research Inst., San Francisco Edward H. Shortliffe, M.D.,Ph.D. Stanford University Johnson & Johnson 1 year $50,000 (*)	0.65	61.20	303
5) SECS Project "Simulation & Evaluation of Chemical Synthesis" W. Todd Wipke, Ph.D. U. California, Santa Cruz NIHEHS ES02845-02 4/82-3/85 $257,801 4/84-3/85 $89,140 Evans & Sutherland Corp. Equipment gift Value $95,000 Stauffer Chemical Co. $6,000	264.61	9877.34	10500
6) SOLVER Project "Problem Solving Expertise" Paul E. Johnson, Ph.D. William B. Thompson, Ph.D. Control Data Corp. (Johnson) 1983-85 $90,000 Microelect. and Info. Ctr. Univ. of MN (Plus Two Colleagues) 1984-1987 $800,000	5.76	356.23	492

7) *** [Rutgers-AIM] *** Rutgers Research Resource "Computers in Biomedicine" Saul Amarel, D.Sc. Casimir Kulikowski, Ph.D. Sholom Weiss, Ph.D Rutgers U., New Brunswick NIH RR-00643-12 (Amarel, Kulikowski) 12/82-11/83 $405,304 NIH RR-02230-01 (Kulikowski, Weiss) 12/83-11/87 $3,198,075 12/83-11/84 $989,276	0.52	38.59	1117
8) AIM Pilot Projects	65.85	2227.48	2461
9) AIM Administration	.93	118.75	686
10) AIM Users	57.36	3836.19	9649
	----------	----------	-------
Community Totals	455.56	17654.52	35676

Stanford Community	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) AGE Project (Core) "Attempt to Generalize" Edward A. Feigenbaum, Ph.D. Dept. Computer Science ARPA MDA903-80-C-0107 (***) (partial support)	11.80	845.30	4076
2) AI Handbook Project (Core) Edward A. Feigenbaum, Ph.D. Dept. Computer Science ARPA MDA903-80-C-0107 (**) (partial support)	11.03	980.94	4425
3) DENDRAL Project "Resource Related Research: Computers in Chemistry" Carl Djerassi, Ph.D. Dennis H. Smith, Ph.D. Dept. Chemistry NIH RR-00612-13 5/82-4/83 $170,710	3.72	183.81	2980
4) EXPEX Project "Expert Explanation" Edward H. Shortliffe, M.D.,Ph.D. Dept. Medicine ONR NR 049-479 1/81-12/83 $456,622 ONR NR049-479 Michael Genesereth 1/84-12/86 $312,070 NSF IST83-12148 Bruce G. Buchanan 3/84-2/87 $330,000 (*) 3/84-2/85 $99,410 (*)	53.75	2391.40	4920
5) GUIDON-NEOMYCIN Project "Exploration of Tutoring & Problem-solving Strategies" Bruce G. Buchanan, Ph.D. William J. Clancey, Ph.D. Dept. Computer Science ONR/ARI N00014-79-C-0302 3/79-3/85 $683,892	45.44	4418.68	5967

6) MOLGEN Project "Applications of Artificial Intelligence to Molecular Biology" Edward A. Feigenbaum, Ph.D. Peter Friedland, Ph.D. Charles Yanofsky, Ph.D. Depts. Computer Science/ Biology NSF MCS-8310236 (Feigenbaum, Yanofsky) 11/83-10/84 $139,215 (*)	106.92	7734.34	10448
7) ONCOCIN Project "Knowledge Engineering for Med. Consultation" Edward H. Shortliffe, M.D.,Ph.D. Dept. Medicine NLM LM-03395 (Shortliffe/ONCOCIN) Edward A. Feigenbaum, Ph.D. 7/79-6/84 $497,420 7/83-6/84 $95,424 NLM LM-00048 7/79-6/84 $196,425 7/83-6/84 $39,502 ONR NR 049-479 1/81-12/83 $456,622 (*) NIH RR-01613 7/83-6/86 $624,455 7/83-6/84 $220,371 NLM LM-04136 8/83-7/86 $211,851 8/83-7/84 $60,517 H.J. Kaiser Family Fdn. 7/83-6/86 $150,000 7/83-6/84 $50,000 ONR N00014-81-K-0004 Michael R. Genesereth (Shortliffe) 1/84-12/86 $512,070 (*) NSF IST83-12148 Bruce G. Buchanan (Shortliffe) 3/84-2/87 $330,000 (*) 3/84-2/85 $99,410 (*)	239.97	14404.62	14389
8) PROTEIN Project "Heuristic Comp. Applied to Prot. Crystallog." Edward A. Feigenbaum, Ph.D. Dept. Computer Science NSF MCS-81-17330 1/82-1/83 $28,976	4.79	635.43	1296

	CPU (Hours)	Connect (Hours)	File Space (Pages)
9) RADIX Project "Deriving Medical Knowledge from Time-Oriented Clinical Databases" Robert L. Blum, M.D. Gio C.M. Wiederhold, Ph.D. Depts. Computer Science/ Electrical Engrg. NSF IST-8317858 (Blum) 3/84-3/86 $89597 (*) NLM (Wiederhold) 5/84-11/86 $291,192	79.44	3140.27	8777
10) Stanford Pilot Projects	61.55	4115.02	6097
11) HPP Core AI Research	383.07	29073.96	42202
12) HPP Associates	57.37	1600.31	2997
13) Stanford Associates	27.01	1016.59	1681
14) Medical Information Sciences	5.62	1315.64	587
	---------	----------	-------
Community Totals	1091.46	71856.29	110842

SUMEX Staff	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Staff	288.21	17591.82	23292
2) System Associates	16.65	1983.43	7847
	---------	----------	-------
Community Totals	304.87	19575.25	31139

System Operations	CPU (Hours)	Connect (Hours)	File Space (Pages)
1) Operations	530.54	67375.43	167863
	=====	=====	=====
Resource Totals	2382.43	176461.50	345520

(*) Award includes indirect costs.

(**) Supported by a larger ARPA contract MDA-903-80-C-0107 awarded to the Stanford Computer Science Department:

System Reliability

System reliability for the DECsystem 2060 has been much better than with our previous KI10 system. We have had very few periods of particular hardware or software problems. The data below covers the entire period in which the SUMEX-AIM community has used the 2060. The actual downtime was rounded to the nearest hour.

7	18	1
Feb	Mar	Apr

Table 1 : System Downtime Hours per Month - February 83 through Apr 83

11	11	1	2	6	0	11	15	26	13	16	28
May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr

Table 2 : System Downtime Hours per Month - May 83 through Apr 84

Reporting period	:	462 days, 23 hours, 41 minutes, and 42 seconds
Total Up Time	:	454 days, 5 hours, 16 minutes, and 57 seconds
PM Downtime	:	1 days, 14 hours, 2 minutes, and 55 seconds
Actual Downtime	:	7 days, 4 hours, 21 minutes, and 50 seconds
Total Downtime	:	8 days, 18 hours, 24 minutes, and 45 seconds
Mtbf	:	2 days, 16 hours, 30 minutes, and 16 seconds
Uptime Percentage	:	98.45

Network Usage Statistics

The plots in Figure 13 and Figure 14 show the monthly network terminal connect time for the TYMNET and the INTERNET usage. The INTERNET is a broader term for what was previously referred to as Arpanet usage. Since many vendors now support the INTERNET protocols (IP/TCP) in addition to the Arpanet, which converted to IP/TCP in January of 1983, it is no longer possible to distinguish between Arpanet usage and Internet usage on our 2060 system.

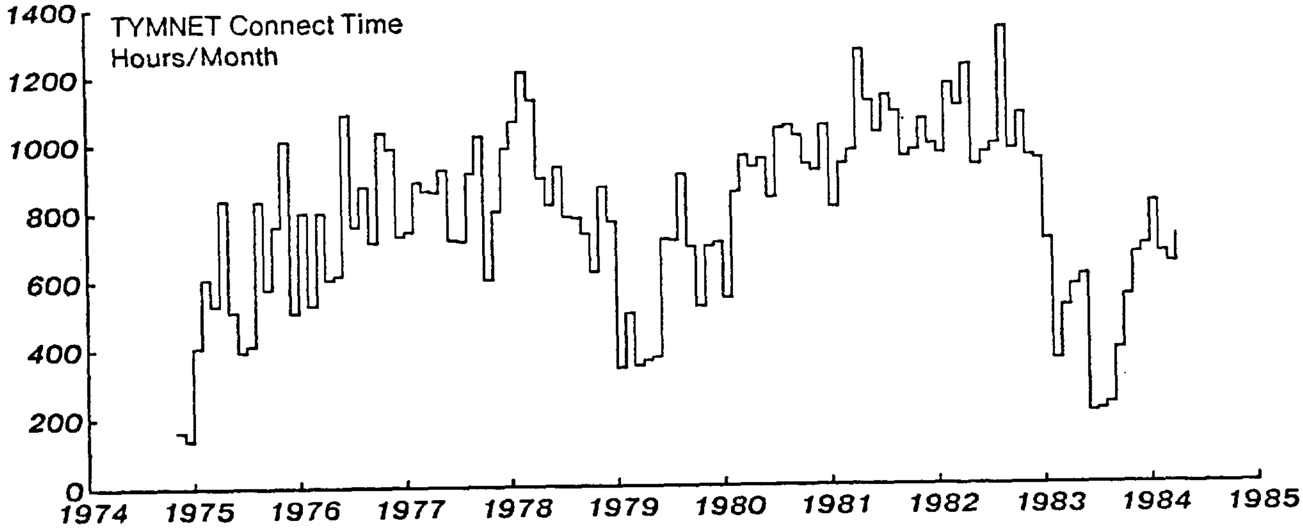

Figure 13: TYMNET Terminal Connect Time

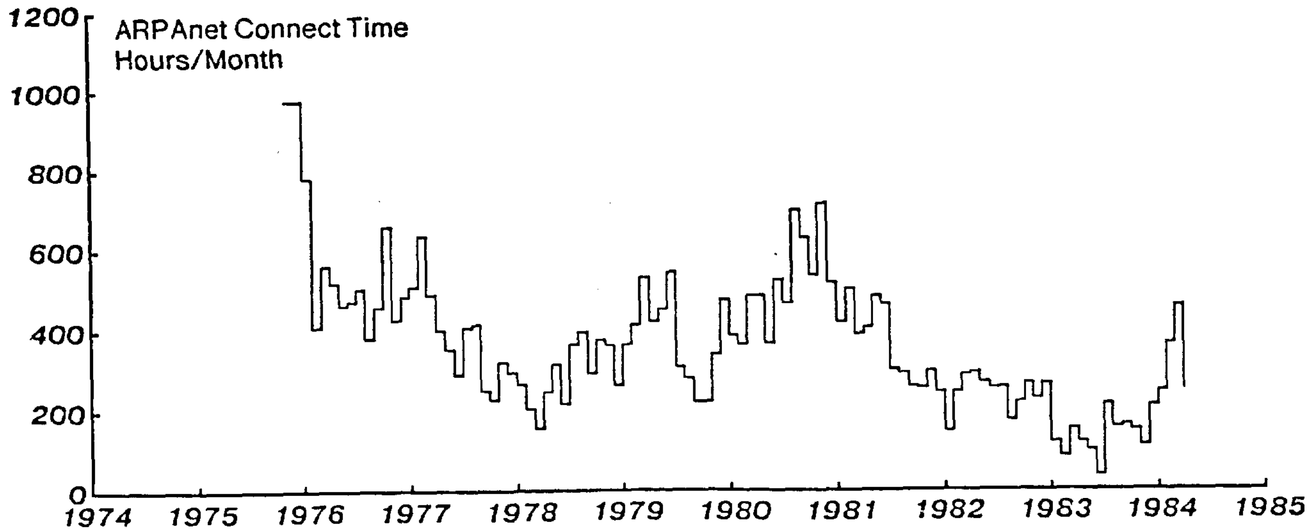

Figure 14: ARPANET Terminal Connect Time

I.A.2.8. SUMEX Staff Publications

The following are publications for the SUMEX staff and include papers describing the SUMEX-AIM resource and on-going research as well as documentation of system and program developments. Many of the publications documenting SUMEX-AIM community research are from the individual collaborative projects and are detailed in their respective reports (see Section II on page 69). Publications for the AGE and AI Handbook core research projects are given there.

1. Carhart, R.E., Johnson, S.M., Smith, D.H., Buchanan, B.G., Dromey, R.G., and Lederberg, J., *Networking and a Collaborative Research Community: A Case Study Using the DENDRAL Programs.* IN P. Lykos (Ed.), COMPUTER NETWORKING AND CHEMISTRY, ACS Symposium Series, No. 19, 1975.

2. Levinthal, E.C., Carhart, R.E., Johnson, S.M., and Lederberg, J.: *When Computers Talk to Computers.* Industrial Research, November, 1975.

3. Wilcox, C.R., *MAINSAIL - A Machine-Independent Programming System.* Proc. DECUS Symposium 2(4), Spring, 1976.

4. Wilcox, C.R.: *The MAINSAIL Project: Developing Tools for Software Portability.* Proc. SCAMC, October, 1977, pp. 76-83.

5. Lederberg, J.L.: *Digital Communications and the Conduct of Science: The New Literacy.* Proc. IEEE 66(11), November, 1978.

6. Wilcox, C.R., Jirak, G.A., and Dageforde, M.L.: *MAINSAIL - Language Manual.* Stanford University Computer Science Report STAN-CS-80-791, 1980.

7. Wilcox, C.R., Jirak, G.A., and Dageforde, M.L.: *MAINSAIL - Implementation Overview.* Stanford University Computer Science Report STAN-CS-80-792, 1980.

In addition, a substantial continuing effort has gone into developing, upgrading, and extending documentation about the SUMEX-AIM resource. These efforts include user guides, help files, and introductory notes, an ARPANET Resource Handbook entry, and policy guidelines.

I.A.2.9. Future Plans

Our plans for the next grant year are based on the Council-approved plans for our 5-year renewal that began in August, 1980. In addition to the specific plans for the next grant year, we include a summary of the overall objectives for this 5-year period to serve as a background. Near- and long-term objectives and plans for individual collaborative projects are discussed in Section II beginning on page 69.

Overall Goals

The goals of the SUMEX-AIM resource are long-term in supporting basic research in artificial intelligence, applying these techniques to a broad range of biomedical problems, experimenting with communication technologies to promote scientific interchange, and developing better tools and facilities to carry on this research. Just as the tone of our renewal proposal derives from the continuing long-term research objectives of the SUMEX-AIM community, our approach derives from the methods and philosophy already established for the resource. We will continue to develop useful knowledge-based software tools for biomedical research based on innovative, yet accessible computing technologies.

For us it is important to make systems that work and are exportable. Hence, our approach is to integrate available state-of-the-art hardware technology as a basis for the underlying software research and development necessary to support the AI work. SUMEX-AIM will retain its broad community orientation in choosing and implementing its resources. We will draw upon the expertise of on-going research efforts where possible and build on these where extensions or innovations are necessary. This orientation has proved to be an effective way to build the current facility and community.

We have built ties to a broad computer science community; have brought the results of their work to the AIM users; and have exported results of our own work. This broader community is particularly active in developing technological tools in the form of new machine architectures, language support, and interactive modalities.

Toward a More Distributed Resource

The initial model for SUMEX as a centralized resource was based on the high cost of powerful computing facilities, which were not readily duplicated. This role is evolving, though, with the introduction of more compact and inexpensive computing technology. Our future goals are guided by community needs for more computing capacity and improved tools to build more effective expert systems, and to test operational versions of AI programs in real-world settings. In order to meet these needs, we must take advantage of a range of newly-developing machine architectures and systems. As a result, SUMEX-AIM will become a more distributed community resource with heterogeneous computing facilities tethered to each other through communications media. Many of these machines will be located physically near the projects or biomedical scientists using them.

The Continuing Role of SUMEX-Central

Even with more distributed computing resources, the central resource will continue to play an important role as a communications crossroad, as a research group devoted to integrating the new software and hardware technologies to meet the needs of medical AI applications, as a spawning ground for new application projects, and as a base for local AI projects. A key challenge will be to maintain the scientific community ties that grew naturally out of the previous colocation within a central facility.

Summary of Five-year Objectives

The long-term objectives of the SUMEX-AIM resource nucleus during the follow-on 5 year period (of which we are in the third year) are summarized below. These are broken into three categories: ***resource operations***, ***training and education***, and ***core research***.

Resource Operations

1. Maintain the vitality of the AIM community -- We will continue to encourage and explore new applications of AI to biomedical research and improve mechanisms for inter- and intra-group collaborations and communications. While AI is our defining theme, we may entertain exceptional applications justified by some other unique feature of SUMEX-AIM essential for important biomedical research. To minimize administrative barriers to the community-oriented goals of SUMEX-AIM and to direct our resources toward purely scientific goals, we plan to retain the current user funding arrangements for projects working on SUMEX facilities. User projects will fund their own manpower and local needs; will actively contribute their special expertise to the SUMEX-AIM community; and will receive an allocation of computing resources under the control of the AIM management committees. There will be no "fee for service" charges for community members. We also will continue to exploit community expertise and sharing in software development, and to facilitate more effective information-sharing among projects.

2. Provide effective computational support for AIM community goals -- We will continue to expand support for artificial intelligence research and new applications work, to develop new computational tools to support more mature projects, and to facilitate testing and research dissemination of nearly operational programs. We will continue to operate and develop the existing central facility as the nucleus of the resource. We will acquire additional equipment to meet developing community needs for more capacity, larger program address spaces, and improved interactive facilities. New computing hardware technologies becoming available now and in the next few years will play a key role in these developments, and we expect to take the lead in this community for adapting these new tools to biomedical AI needs.

3. Provide effective and geographically accessible communication facilities to the SUMEX-AIM community for effective remote collaborations, communications among distributed computing nodes, and experimental testing of AI programs -- We will retain the current ARPANET and TYMNET connections for at least the near-term and will actively explore other advantageous connections to new communications networks and to dedicated links.

Training and Education

1. Assist new and established projects in the effective use of the SUMEX-AIM resource -- Collaborative projects continue to be responsible for the development and dissemination of their own AI programs, but the resource staff will provide general support and will work to make resource goals and AI systems known and available to appropriate biomedical scientists. We will continue to exploit particular areas of expertise within the community for developing pilot efforts in new application areas.

2. Continue to allocate "collaborative linkage" funds to qualifying new and pilot projects to provide for communications and terminal support pending formal approval and funding of their projects -- These funds are allocated in cooperation with the AIM Executive Committee reviews of prospective user projects.

3. Continue to support workshop activities including collaboration with the Rutgers Computers in Biomedicine resource on the AIM Community Workshop and with individual projects for more specialized workshops covering specific application areas or program dissemination

Core Research

1. Continue to explore basic Artificial Intelligence research issues for knowledge acquisition, representation, and utilization; reasoning in the presence of uncertainty; strategy planning; and explanations of reasoning pathways with particular emphasis on biomedical applications -- SUMEX core research funding is complementary to similar funding from other agencies and contributes to the long-standing interdisciplinary effort at Stanford in basic AI research and expert system design. We expect this work to provide the foundation for increasingly effective consultative programs in medicine and for more practical adaptations of this work within emerging microelectronic technologies.

2. Support community efforts to organize and generalize AI tools that have been developed in the context of individual application projects -- This will include work to organize the present state-of-the-art in AI techniques through the development of practical software packages for the acquisition, representation, and utilization of knowledge in AI programs. The objective is to evolve a body of software tools that can be used to more easily build future knowledge-based systems and explore other bionmedical AI applications.

Specific Plans for Year 12

Specific plans for the next grant year (12) are summarized in the paragraphs below. The directions and background for much of this work were given in earlier progress report sections and are not repeated in detail here.

Professional Workstations

We see our major development efforts in year 12 to be in the area of professional workstations, and specifically, to fine tune the integration of these workstations into our networking environment. This involves software integration, support of network protocols, general access to network printing facilities, telnet access to Lisp machines, and overall workstation maintenance and support.

We will also continue to explore the use of low cost workstations within our environment, both as distributed processors for text editing and electronic mail, and as powerful graphic terminals for use with sophisticated programs running on our mainframes. We also see the use of virtual graphics interfaces running on remote workstations to be of continued importance to our progress in the future.

Continued Operation of Existing Hardware

The current SUMEX-AIM facilities represent a large existing investment. We plan to continue development of our main timesharing machine, the DEC2060/TOPS-20 system, and the SUMEX-AIM file server (SAFE), and make changes as necessary to improve the performance of these machines. We do not propose any substantial changes to the other hardware systems (2020, shared VAX, and Lisp Machines). We expect them to continue to provide effective community support and serve as a nucleus for our distributed resource.

Communication Networks

Networks have been centrally important to the research goals of SUMEX-AIM and will become more so in the context of increasingly distributed computing. Communication will be crucial to maintain community scientific contacts, to facilitate shared system and software maintenance based on regional expertise, to allow necessary information flow and access at all levels, and to meet the technical requirements of shared equipment.

We have had reasonable success at meeting the geographical needs of the community during the early phases of SUMEX-AIM through our ARPANET and TYMNET connections. These have allowed users from many locations within the United States and abroad to gain terminal access to the AIM resources and through ARPANET links to communicate much more voluminous file information. Since many of our users do not have ARPANET access privileges for technical or administrative reasons, a key problem impeding remote use has been the limited communications facilities (speed, file transfer, and terminal handling) offered currently by commercial networks. Commercial improvements are slow in coming but may be expected to solve the file transfer problem in the next few years. A number of vendors (AT&T, IBM, XEROX, etc.) have yet to announce commercially-available facilities, but TELENET is actively working in this direction. We plan to continue experimenting with improved facilities as offered by commercial or government sources in the next grant term. We have budgeted for continued TYMNET service and an additional amount annually for experimental network connections.

High-speed interactive terminal support will continue to be a problem since one cannot expect to serve 1200 to 9600 baud terminals effectively over shared long-distance trunk lines with gross capacities of only 9600 to 19200 baud. We feel this is a problem