Academic Computing and Networking at NYU is edited and published by New York University's Academic Computing Facility (ACF). Its scope includes information about computing and networking activities at NYU's various schools, departments, and administrative units, and outside developments of interest to the NYU community.

Copies of Academic Computing and Networking at NYU are mailed to university faculty and staff, and are also available from the ACF's Information Services Office (room 312, Warren Weaver Hall), at the ACF Help Center (second floor, Warren Weaver Hall), and at the ACF computer labs (listed inside the back cover). Students holding ACF individual computer accounts are included automatically in the newsletter's mailing list.

Selected articles from this publication have been made available on the NYU CWIS, starting with the March 1993 issue. To locate these articles, choose Academic Computing and Networking Resources from the main CWIS menu, then select Academic Computing Facility, then Publications and Reference Collections, and finally, Academic Computing and Networking at NYU (the newsletter).

We welcome your comments and suggestions about the articles in this issue, and about articles for future issues. Contributions are invited for consideration by the editor; please call 998-3038 or send e-mail to frederickson@nyu.edu for more information. Articles are written by members of the ACF staff, unless otherwise indicated.

Opinions expressed in the articles in this publication are those of the authors and not necessarily those of the Academic Computing Facility or of New York University.

Below many of the authors' bylines are electronic mail (e-mail) addresses. If you do not use e-mail but would like to, call the ACF HelpLine at 998-3333 for information about opening an appropriate account.

This issue was prepared on Apple Macintosh Quadra and IiCi computers, using Aldus PageMaker, Microsoft Word, Adobe Type Manager, Adobe Photoshop, and HiJaak, among other programs. Fonts used in this issue are Palatino for the text and Gill Sans Bold for headlines, along with Zapf Dingbats and Courier for special effects; the logo is set in Adobe Garamond bold italic. Camera-ready copy of text and screen shots was produced using a 600-dpi Hewlett-Packard 4si printer. Echo Graphics printed and bound the publication.

Editor
David Frederickson
frederickson@nyu.edu

Production Assistants
Franca Gaudio,
Vijay Jodha, Frank Taylor

Design
David Frederickson

Contributors
David Ackerman, Lisa Barnett,
Gary Chapman, Gregg Geller
(GSAS), Estelle Hochberg,
Jeffrey Lane, Bryan E. Leonard
(Computer Store), Frank LoPresti,
Tom McNulty (Bobst), Richard
Maisel (FAS), Regina R. Monaco
(FAS), Tim O'Connor,
Caroline Hodges Persell (FAS),
Robert A. Yaffee
Thanks also to
Vincent Doogan, Eray Ekici,
Edward A. Friedman,
Tim O'Connor, Lu Ratunil,
Marvin Rich (GSAS), George
Sadowsky, Estarose Wolfson,
Carmen Vásquez

copyright © 1994 New York University
Dilemmas and Pitfalls in Developing Instructional Software: The ISEE Story (1)

Caroline Hodges Persell and Richard Maisel
persell@acfcluster.nyu.edu • maisel@socgate.nyu.edu

(In an NYU colloquium last year, reviewed in the January issue of this publication, Professors Persell and Maisel introduced ISEE [Introduction to Sampling Error Experiments], the computer program they had developed to teach the principles of scientific sampling. Later, they mused on the process of developing the software. In this issue, they discuss the dilemmas facing a person intending to write an instructional program; in the next issue, they will discuss the conditions that impede and facilitate the process.)

Developing instructional software in sociology shares some features with other types of pedagogical innovation, while raising additional issues that are unique. Here we offer a case study of one instance of instructional software development. We focus on the occupational, developmental, technical, and pedagogical dilemmas encountered.

The Instructional Problem

Scientific sampling is one of the most powerful tools at the disposal of social researchers. Principles of sampling are a critical part of undergraduate and graduate education in the social sciences and in many applied fields. The power of sampling is not always fully grasped by students, however. Because of their often limited backgrounds in mathematics, many have serious difficulties learning the basic concepts underlying statistical inference and reasoning. After many years of teaching we decided to collaborate on developing instructional software to address this problem.

In the course of developing this software, we encountered a series of dilemmas.

Occupational Dilemmas

First, developing software poses a series of occupational dilemmas. Three characteristics of the occupation of academic sociologist have significant implications for the development of instructional software. Thoughtful scholars have reflected on the sociological ambivalence both toward teaching and toward computing. To these we can add a third, the tendency to marginalize the development of instructional software. Research brings rewards, recognition, and control that teaching does not. Thus the ambivalence about improving pedagogical practice is not unique to software development; it is shared with other efforts to enhance the teaching of sociology.

Software development faces a unique problem as well: computing is perceived as menial work requiring little skill. The combination of ambivalence toward teaching and ambivalence toward computing is not likely to enhance the status of instructional software development.

A third characteristic is a tendency, shared by both the home discipline (sociology) and by computer science departments, to marginalize software develop-
ment. The former bring the ambivalence noted above and consider the activity "not really sociology," while technical people see it as boring, applied work, offering nothing new or exciting. Thus the developers are left with no intellectual home for their work.

These three characteristics are particularly important in relation to the management of an academic career. People developing educational software without tenure place themselves at great risk. Both the complexity and the lengthy time associated with developing instructional software pose problems in relation to evaluating faculty for promotion and tenure. It is difficult to judge the originality of new software; it generally lacks the peer review that journal articles get before publication and that books receive after publication. Awards given by EDUCOM and Zenith are two efforts to address the issue of quality assessment, but those awards are not usually made by disciplinary peers.

**Developmental Dilemmas**

In this project, there was a clear need for institutional support. We couldn't do the project without outside funding, and that required institutional support. The cost of that support, though, was losing ownership of the product. We agreed that NYU would own the copyright to the software, in return for the various forms of institutional support provided.

A second developmental dilemma was the tension between the imperative of continuous iterative development and the need for systematic evaluation. If software is to be any good, it needs to be tried constantly with users and learners. The lessons learned from those trials need to be fed back into programming and pedagogical improvements. We

---

ISEE on the screen: Above small windows for the menu and commands, a larger window in the Sample Design module (above left) produces different sorts of samples of a population and calculates the results. Students can select as many different samples — simple random, systematic, replicated, stratified, and thin-zone — as they wish. They can see (hence the name ISEE) the patterns produced by various selection procedures and see whether one or another type of sample produces larger or smaller differences between the sample mean and the population mean.

The middle screen shows a histogram for a population with a bimodal distribution, which permits students to draw a series of samples from such a population. When they do so, the sampling distribution builds dynamically before their eyes, and they discover that the distribution of the sample means begins to resemble a normal curve, despite the bimodal distribution in the original population.

The bottom screen shows how students can conduct sampling experiments under as many as four different sets of conditions. In this way they can readily compare the effects on sampling distributions of populations with different characteristics, as well as the effects of different sample sizes and different methods of sampling.
did extensive trials here at NYU and at Queensboro Community College, which provided a great deal of feedback.

The dilemma of the “moving target” was highlighted when we had someone else — Ganka Dimitrova, a graduate student at NYU — do an experimental evaluation of the program. The carefully designed experiments revealed no difference in tested achievement for students who used the program and those who did not. However, by the time the experiments were conducted, we had added many new features to the program, so the testing was to some extent incomplete. Evaluation always lagged behind the current version of the program; by the time we finished version 3.1, we were out of time and money for additional systematic evaluations.

**Technical Dilemmas**

In developing software, there is a tension between ease of programming, and the speed and versatility of the finished product. At first, we opted for ease, developing the preliminary version on a VAX minicomputer running MINITAB, a standard statistical program for data analysis. In retrospect, it’s clear that this was not the most valuable route to follow. Even when translated to run on a PC, the program based on MINITAB worked slowly through the network server. And it didn’t feel like an interactive program: it was static rather than dynamic, and the graphics were feeble.

Another technical dilemma concerned the nature of the equipment required to run the program. Should the goal be creating a program that was universally usable by any instructor, student, or school with at least a PC-XT, as is still the situation in many underendowed colleges? Or, should the program be written for “high-end” users, at that time those with 386 machines, math coprocessors, and color VGA monitors? The dilemma is one of transferability vs. functionality.

Our initial decision favored transferability, since that was one of the primary goals of FIPSE (the Fund for the Improvement of Postsecondary Education), which was funding the project. Hence the first PC version was written for use on the widest range of machines. Our ultimate resolution was to create two versions of the program: ISEE, for PC-XT or AT machines with no color or graphics board, and ISEEG, for machines with color VGA boards. But with two versions, it is difficult to support and improve the nongraphics one. We have since dropped the two versions and have one graphics version called ISEE.

The more advanced system permitted us to solve a pedagogical problem. An understanding of sampling requires one to keep three different types of information in mind at the same time:

- information about the population
- information about how the sample was selected
- characteristics of the sample selected.

With the nongraphics version, students had to keep going back and forth in the program, which disrupted the learning process. We realized that this problem could be solved by changes in the program itself, by putting all three on the screen at the same time in separate windows — a much better solution. We were reluctant to assume that others had such equipment available. But, if our goal is to improve instruction, we should keep a longer time frame in mind and utilize equipment that increases student interest and (hopefully) learning as well.

Another technical dilemma is complexity vs. ease of use in the program itself. The more a program can do, the greater its complexity. As we saw more things we wanted the program to do, the more complex it became. The tension is, can it be kept user-friendly while adding capability?

Finally, there is a technical dilemma between offering an open-ended system that can deal with any new populations a user might add vs. having a program that draws attractive histograms and other graphics. The simple sample module of ISEE resolves this dilemma in favor of a small, closed system, which allows for very tidy histograms. The ultimate goal for ISEE is a system that can handle whatever people need to do with it. Obviously, of course, that is a much greater programming challenge.

**Pedagogical Dilemmas**

A classic pedagogical dilemma exists between spending more time on one topic to achieve mastery vs. covering more topics quickly. We began to realize that ISEE could be used in two quite different ways.

The first was as a simple adjunct to a course as currently taught — perhaps to illustrate points in a lecture or for homework assignments. However, this might not be enough to give students a grasp of the basic ideas underlying sampling theory.

The second way of using the program requires a fundamental change in teaching. Rather than leading the students through the learning, an instructor us-

(continued on page 10)
Electronic CAS: Students Find Info, Virtual Advisor, and Cyberdean Online

Gregg Geller
gellerg@acf2.nyu.edu

“The College of Arts and Science has entered the electronic age,” Matthew Santirocco, the new dean of the College of Arts and Science, announced last month. “With CASIS, Cyberdean, and the CAS Virtual Advisor all online, we are on the way to becoming the Electronic College.”

While there is no substitute for direct contact between the Dean’s Office, the Advising Center, and individual CAS students, the new online services will give students the opportunity to have 24-hour direct communication with the College of Arts and Science. These new services are meant to complement the support network already in place, not to replace it. As outreach and accessibility increase, students will be provided with more efficient service, but they will also feel more closely connected to the College.

CASIS: The College of Arts and Science Information Services

CASIS — the College of Arts and Science Information Service — is an online menu system, part of the NYU CWIS, that offers information of particular interest to CAS students. It grew out of the desire to reach the diverse student body of the College. While full-time students are generally involved in events around Washington Square, part-time and nontraditional students may not always respond to posted flyers and mass-mailings.

Through CASIS, found under Academic Departments and Programs (see the illustrations on the next page) on the NYU CWIS, the College hopes to involve as many students as possible in its educational and extracurricular activities. Since the NYU CWIS is accessible over the Internet, even CAS students who are studying abroad will be able stay current with college events.

CASIS posts up-to-date descriptions of the services offered by the College of Arts and Science. These include the Academic Achievement Program, Career Assistance, Freshman Programs, Engineering, International Programs, Preprofessional Advising, and Tutoring.

In addition, students will soon be able to view the entire CAS Bulletin online, browse current internship listings, dip into the new CAS Senior Handbook, and check a weekly calendar for upcoming workshops, lectures, and events. Students can rely on CASIS to bring them helpful information in a concise and useful format.

CAS Cyberdean

Cyberdean is Dean Santirocco’s alias on NYU e-mail. Student outreach is very important to the dean: “Whether we meet face-to-face or we interface through Cyberdean, students can reach me all day, at night, and even on weekends,” Santirocco said.

Students and faculty alike have responded to Dean Santirocco’s initiative. So far, Cyberdean has received numerous messages from locales ranging from Texas
to Taiwan. Messages to Cyberdean have included requests for information on NYU programs, suggestions for the College, and welcomes for the new dean. In addition to the dean’s walk-in hour (Wednesdays from 4:00 to 5:00, in room 910, Main), electronic mail is another way to ensure that everyone, no matter where, has personal and direct access to the dean of the College.

If you want to send e-mail to Dean Santirocco, address it to cyberdean@is.nyu.edu.

Advisement Online

Although the CAS Advising Center, on the ninth floor of the Main Building, always has an open-door policy, student questions are not always limited to business hours. Even questions that do come up during the day should not always warrant a trip to the Advising Center. The new Virtual Advisor allows CAS students to ask questions at any time of the day or night and get a reply by the next business day. Answers to questions that will benefit other students will be posted on CASIS under the CAS Virtual Advisor menu.

“The CAS Virtual Advisor cannot replace in-person academic advising by faculty or professional advisors,” said Willie Long, Assistant Dean for Advisement. “Instead it expands contact with faculty and staff.” Virtual Advisor will always make referrals to appropriate individuals within the College and University, and every response will end with an invitation to come to the Advising Center to meet with a professional advisor. So, for example, if a student writes to ask the Virtual Advisor about tutoring in mathematics, the response might read:

Math tutors are available through the CAS Tutoring Program, located in 905 Main Building. For more information, see “Tutoring Program” under “CAS Programs” in CASIS. You may also stop by to set up an appointment or call 998-8130 to learn more. The Department of Mathematics also offers math tutoring. You should go to Room 705, Warren Weaver Hall, to pick up a current tutoring schedule. For more information call 998-3005. At the CAS Advising Center, we also have study skills and time management handouts available. Come in and see any of the professional staff advisors for copies.

You may send questions and comments to the CAS Virtual Advisor at CasVirtualAdvisor@nyu.edu.

The Future of the Electronic College

“This is only the beginning,” Dean Santirocco said. “Our online programs are all part of a larger vision. These services will empower students and ultimately transform the College from a teaching to a learning environment.” Electronic mail, for instance, is increasingly being recognized as a teaching tool. Some professors in the College are already accessible to students through electronic mail, and they could soon require students to sign up for electronic newsgroups as an extension of classroom dialogue, and to access online curricula and other course materials. This type of learning environment will change the way College of Arts and Science classes are taught and the way students learn. “It is all very exciting,” Dean Santirocco added. “We are on the verge of applying tomorrow’s interactive technologies to today’s educational processes.”

Stay tuned!
Now that the NYU CWIS is well into its second year, it has established itself as a central resource for information about all facets of New York University, available both on campus and around the world. The CWIS — NYU’s Campus-Wide Information System, developed and operated by the Academic Computing Facility, with content provided by the various schools and departments of the university — has experienced a tremendous growth both in quantity of information and in its utilization since it was launched in September of 1993.

More Info, More Calls

The number of accesses — each time a user asks for a new menu item counts as an access — grew from the outset, and during the fall 1993 semester it peaked at about 10,000 accesses per week. During the spring, usage peaked in April at about 17,000 per week, up 70 percent from the fall’s peak. During the summer, surprisingly, usage continued to increase to about 21,000 per week, and in mid-fall 1994, it has grown to about 40,000 — doubling and redoubling within the first year. On the Internet, traffic from the use of Gopher — the program behind the NYU CWIS and thousands of similar information servers — is increasing at a rate of 10 to 20 percent per month.

The number of offices providing information on the NYU CWIS is increasing at a similar pace. At the end of 1993 there were perhaps a dozen information providers. Now, more than three dozen departments have information in the system, and more than a dozen more are planning additions soon.

There are now seven Gopher servers (the computers that store the information, using Gopher server software) at NYU — and as of November 8, there were 5,323 public Gopher servers around the world. The ACF now supports six Gopher clients (the programs you use, either on your desktop computer or through a shared ACF computer, to reach the CWIS): TurboGopher for the Mac, PC Gopher, WSGopher (a Windows program), and those for the multiuser operating systems UNIX, VMS, and VM.

Two New Searching Tools

The worldwide growth in the number of Gopher servers and in the quantity of information has magnified the need for better searching mechanisms. A new suite of tools to search the NYU CWIS was made available this fall, which should make finding information easier than ever.

The CWIS now has two methods of searching built in. Every day, in the early hours of the morning, two...
search programs look through all the items on the CWIS and catalog it, thus making sure that your next search will be up-to-date. These two programs are called WAIS (pronounced waysz) and Jughead. To use these two search programs, you need only select Search NYU CWIS from the main menu.

**Searching Menu Titles with Jughead**

If you then select Search Menu Titles, you will launch Jughead, a program that will search
- for the word or words (whole or partial) you want
- only in the titles (not in the text itself)
- of all the menu items (both directories and documents)
- in all the Gopher servers at NYU (not just the main NYU CWIS server).

You will be prompted for a search string. You should fill in the word or words that you want to search for. The search is Boolean; that is, you can specify more than one word for the search, with the words and, or, and not used as logical operators. This simply means that searching for New and York will return all the NYU CWIS entries with both the words New and York in the menu title. If no operator is specified, the word and is assumed between words. So, searching for New York is the same as searching for New and York. A search for New not York would return titles such as New and Noteworthy and New Jersey but not New York University. Capitalization of the letters is ignored: Jughead thinks York is the same as YORK and York.

You can also use an asterisk (*) in the search string, as a substitute for unspecified (or absent) letters, but only as the last character of any word in the string, not as the first or only character, and not if it is followed by any other characters. So, a search for new* returns titles with the words new, newspaper, newsgroup, and so on.

Sometimes a search will find multiple occurrences of the same title. In fact, these titles occur on several different menus within the NYU CWIS. They may lead you to the same information or they may lead to different objects; Jughead has no way of knowing.

Jughead, by the way, is an acronym for Jonzy's Universal Gopher Hierarchy Excavation and Display; the program was written by Rhett “Jonzy” Jones at the University of Utah.

**Searching Documents with WAIS**

If you select Search the Text of All Documents, the computer will run WAIS, which will search
- in the text (not the titles on the menu) of all articles
- on the main NYU CWIS computer only
- for any word or words you specify.

You will be prompted for a search string, the word or words you want to search for. This is not a boolean search; that is, you cannot specify the logical connection between words in your string. The search will give you any article that has any of the words you type, logically as if the word or appears between the words.

For example, if you search for New York, all documents with either the word New or York inside the document will be returned. Note that for a WAIS search, neither New nor York need be in the title of the document.

Only items that reside on the central NYU CWIS Gopher server are included in this search. Thus, items on the Medical Center server would not appear.

WAIS stands for Wide Area Information Servers. Originally, the program was developed by Thinking Machines, Apple Computer, and Dow-Jones.
NETWORK NEWSBYTES

Netscape: New Web Browser
Nearly Ready for Prime Time

Now making waves on the Internet is a new entry called Netscape, a suite of programs for dealing with the Internet through the World-Wide Web protocol.

We've written in several recent issues about the Web and two of the programs for browsing it — Lynx (the character-based program from the University of Kansas that's included with the NYU-Internet account; see the September issue for more) and NCSA Mosaic (the graphical interface from the National Center for Supercomputer Applications that works well only on relatively powerful computers with direct network connections).

A new firm called Netscape Communications Corporation, led by Silicon Graphics founder Jim Clark and Mosaic creator Marc Andreessen, has introduced Netscape. There are three versions of the client software (for the PC, Mac, and Unix platforms); they are now in version 0.9, and are available free at several FTP sites, for those of you who want to try it out on your own computers and help find the bugs in the program. In offering the client software free to individual, academic, and research users, Netscape Communications sees itself as continuing an established Internet tradition and contributing to the explosive growth of new information applications on global networks.

I've tried the Mac version out, and though I've heard of bugs, it seems to work well — so far — and work swiftly. If you've been using Mosaic and have been bothered by the pace at which images download, and you have the requisite sense of adventure, you might give Netscape a try. (If you're cautious about putting prerelease versions of programs on your computer, wait at least until version 1.0 comes out.) Using Lynx or Mosaic, look for the announcement on the Web at http://mcom.com

The White House
Is Part of the Web Now, Too

It's time for another update on the electronic White House. In past issues, we've noted the Clinton administration's efforts to be accessible by e-mail (president@whitehouse.gov) and to make masses of federal information available online (more about that below).

On October 20, the White House unveiled its latest and most sophisticated electronic interface: a World-Wide Web server and an interactive multimedia structure of information, which you can reach at the address http://www.whitehouse.gov and can view on Lynx (without sound and graphics, through your NYU-Internet account) or, with all the bells and whistles, through Mosaic or Netscape (see illustration below and the Newsbyte above).

From the “home page,” you can...
use your mouse to call up photos of the first family, audio welcomes by President Clinton and Vice-President Gore, a tour of the White House, or a guest book in which you can enter your comments and, if you wish, your name and e-mail address. Other options lead to information sources further afield: in the executive branch or other branches of the government, or lists of publications available (over 3,000, by one report). You can view a map of Washington, D.C., enlarge sections of it, and click on buildings such as the Capitol and the Supreme Court to bring you to further information and menu options. Some of these connect to text-only Gopher servers, often housed on computers across the country.

This polished new effort reflects the administration’s oft-stated commitment to the national information infrastructure. The White House Web server is “mirrored” (that is, its contents are maintained on other computers, to which your log-on may be diverted) at two NASA sites — Goddard in Maryland, and Ames in California.

Many government agencies are making information available electronically, and the objective is that in time, all will. Some agencies have their own Gopher servers, but many universities offer federal information on theirs as well, sometimes better organized than the originals. A few Gopher sites to try:

For documents from the Bureau of Justice Statistics: connect via Gopher to uacsc2.albany.edu and look under United Nations Justice Network.

For information on health-care reform, connect via Gopher to ace.esusda.gov and look under Americans Communicating Electronically, then National Policy Issues, and finally Health Care Reform Agenda. Other issues are covered in related menus.

For the full text of all bills in the House of Representatives, try gopher.house.gov; look under Congressional Information, then Legislative Resources.

The real treasure trove, though, is FedWorld, a central online repository of government information maintained since 1992 by the National Technical Information Service (NTIS). FedWorld comprises over 130 bulletin boards not previously available via the Internet; it offers a central access point for locating and acquiring government information. Replete with details on everything from visa requirements for entry into Pago Pago to buying surplus army materials, Fedworld can be reached at fedworld.gov via Telnet.

More Federal Info Online — Finding the Trees in the Forest

Do you think you could use a little guidance on traveling the information superhighway? You wouldn’t be the only one. Just before Labor Day, the Washington Post ran a five-part series of articles called “Government Online,” an overview of the state of the electronic nation, federal style.

Among the lawmakers listed in one entry were nine senators (including from the tristate area, only sendodd@dodd.senate.gov), thirty-six representatives (including boehlert@hr.house.gov from New York, and two from Connecticut: cshays@hr.house.gov and the more opaque bozrah@hr.house.gov for Sam Gejdenson).

The real treasure trove, though, is FedWorld, a central online repository of government information maintained since 1992 by the National Technical Information Service (NTIS). FedWorld comprises over 130 bulletin boards not previously available via the Internet; it offers a central access point for locating and acquiring government information. Replete with details on everything from visa requirements for entry into Pago Pago to buying surplus army materials, Fedworld can be reached at fedworld.gov via Telnet.

The Internet Posts a New Growth Record in Q3, ’94

It’s no longer news that the Internet is growing at a phenomenal rate, though few can say for sure what the rate is, or what the totals are. If you’ve been trying to make sense of it all, you’ll have noted that one article says there are ten million people on the Internet, and the next says twenty. No one knows for sure.

The Internet comprises 45,000 autonomous networks, and as such “defies definitive measurement,” according to Network Wizards (www.nw.com), who regularly try to measure it, nonetheless. Every three months, they do “a complete search of the Domain Name System,” which yields about the only hard figures available, for the number of Internet “hosts” — that is, computers with their own Internet Protocol (IP) addresses — in the world. On July 1, 1994, there were 3.2 million of them; on October 1, there were nearly 3.9 million: a 21 percent increase for the quarter. (Most of those addresses are in North America — 2.7 million. NYU has 6,500 of them — about 0.167 percent.) Each of those addresses potentially represents a computer connected directly to the Internet, although not all are active all the time (some addresses are reserved for future use; some computers are connected only part time).

Those of us in the .edu domain have probably assumed that it will always be the largest, and it does have nearly a million hosts (982,181, to be exact). But it has just been surpassed by the 1,054,522 commercial hosts in the .com domain, up 36 percent in the quarter. There goes the neighborhood, folks.

It’s anybody’s guess how many people a given computer represents: probably at least one, at a minimum, but at the other extreme, a single computer can serve several thousand users, as the machines known as acfcluster and is do at NYU. And what does it mean to be “on the Internet”? Able to send and receive e-mail? To read electronic bulletin boards regularly? To spend two hours a day surfing the Internet,
burrowing in Gopherspace, or spinning the World-Wide Web? Given such a range of numbers and definitions, the best that anyone can do is to make an educated guess, and call it an estimate: "At last count there were an estimated five million people on the Internet," or ten million, or twenty million — and if there were only ten million yesterday, there will surely be twenty tomorrow. The figures are slippery, but large. And they are growing exponentially.

The Names People Play

Enough numbers; how about names? The Network Wizards also offer a list of the fifty most common host names (excluding anything with a number in it; NYU’s is would qualify, but not is2, apparently). Ten or a dozen of the names are strictly utilitarian — www and mail and news and ns, for name server, which, though unimaginative, ranks number 1, appearing on 1,742 hosts. Mythology is well represented, with venus (1,342) in second place and pluto (1,230) in fourth, followed by mars in fifth. If you combined the Greek and Latin names for the next chap (or planet?) — zeus (1,058) and jupiter (1,057) — he’d come out first, but unfortunately he split his ticket and lost out to the lady on the half-shell. Not until number 23 do we reach real people with newton (700); gauss (584) is number 33. Number 43 is homer (541) — but whether that’s the author or the Simpson may never be determined.

It’s not likely that calvin (593) and hobbes (582) are theologian and philosopher, respectively, except at the high allegorical level of the comic strip, whence also snoopy (553) and mickey (536). Still holding up his end of things in fiftieth place is that paragon of mythological reliability (and outsmartability), atlas.

— David Frederickson
frederickson@nyu.edu

ACF HelpLine Q&A

Q: My Internet address is pdq1234@is2.nyu.edu. I understand the nyu part; what does the rest mean?

A: Just like your home address, your e-mail address goes from the most specific to the most general. Your username is pdq1234 — your initials and part of your ID number. Your account is on a machine called is2, which is found at NYU. NYU is an educational institution, so it is in the Internet domain edu. (Other domains in the U.S. are gov for government, com for commercial, org for nonprofit organizations, etc.)

There are several machines here at NYU that support e-mail, so other e-mail addresses will be different from yours.

Student usernames follow the format of abc1234; staff, faculty, and administrative account are some variation of the person’s last name — sometimes minus a vowel or two, or with the first initial added.

Here are some possible e-mail addresses at NYU these days:

- pdq1234@is2.nyu.edu
- abc5678@acfcluster.nyu.edu
- xyz1004@is.nyu.edu
- franklnb@acf2.nyu.edu
- adamsjq@acf4.nyu.edu

— L. Barnett

Call the ACF HelpLine at 998-3333

Dilemmas and Pitfalls (continued from page 3)

ing ISEE can point students in a given direction by asking questions. Students can then, with some help from the instructor, discover answers for themselves by conducting appropriate experiments with the program. As one example, when students in a lab draw samples on their own PCs and begin comparing their results, they quickly notice that they get different answers. This experience sensitizes them to sampling error and motivates them to learn more about it. An active, structured learning procedure is more likely to break through the learning barriers — but it takes more time than is usually scheduled for this topic.

The first approach would be preferable for students with greater mathematical skills, as it would allow them to cover more material. But it would not help students with less mathematical skill, for whom the program was initially designed. The second, slower, approach benefits them, but penalizes better-prepared students by limiting the amount of material that can be covered in the course. ISEE does permit students to spend more time learning conceptually challenging material on their own, without reducing the class time available for other material.

Of course, identifying dilemmas is not enough. Nor can individuals working alone always resolve them. Instead, social conditions can facilitate the development of instructional software; specifically, various kinds of institutional support can play an essential role. In the next issue we will explore these social factors and creative conditions that affect software development.

 Academic Computing and Networking at NYU
NYU Gains High-Speed Connection to the Internet

As this issue goes to press, the connection from NYU-NET to the Internet is being upgraded in speed and in reliability. The present connection, via PSInet, operates at a speed of 1.5 megabits per second; the new connection, via Sprint, will operate at 45 megabits per second — a thirtyfold increase in capacity, or bandwidth. The new connection should be ready in November or December. At the same time, many other institutions in New York State are participating in similar upgrades. Members of the NYU community can therefore expect superb connectivity between NYU and other state resources, and generally better communications to the Internet at large. These improvements derive from the upgrade of NYSERnet, of which NYU is a founding member.

NYSERNet — the New York State Education and Research Network — recently announced agreements with NYNEX and Sprint to upgrade its network and internetworking services for New York State. This arrangement is the first step in NYSERNet’s five-year plan to increase its information-carrying capabilities more than a hundredfold. Already by the end of 1994, NYSERNet and New York State will have the most technologically advanced state data network in the country.

The NYNEX and Sprint agreements call for a network capacity of 45 million bits (45 megabits) per second — known as a T3 network — for NYSERNet affiliates by this month. This will vastly improve routing and service, and deliver T3 speeds for prices that are less than half the current national average. Until now, NYSERNet has operated a T1 network, capable of 1.5 million bits per second.

Because NYSERNet is working with NYNEX and Sprint, “access to the network and Internet will eventually be provided throughout the state through a local dial-up phone call anywhere,” said James Luckett, executive director and vice-president of NYSERNet. “This will guarantee that all New York State citizens have assured and equal access to information resources and the other benefits of the national information infrastructure.”

Sprint will initially provide a 45-megabit high-speed statewide area network, which will include 45-megabit connectivity to the Internet, comprehensive management of the new NYSERNet network, and management of the NYSERNet affiliates’ TCP/IP-based services onto that network and onto the Internet. Sprint will also provide communications and routing equipment; NYNEX will provide the backbone — the actual lines. This network should grow quickly to gigabit speeds over the next few years.

The NYSERNet five-year plan will give New York State the first production gigabit communication network in the U.S. It will be capable of providing high-speed access at all the state’s institutions of higher education, all of its 6,000 K-12 schools, its 7,000 libraries and museums, 500 health-care institutions, 1,500 municipal, county, state, and federal agencies, and some 15,000 commercial users who interact with these sectors.

NYSERNet’s new network will support all traditional Internet services, as well as such new services as interactive two-way video and teleconferencing, and electronic-commerce applications. It will enable a new generation of inexpensive distance learning, telemedicine, and community network applications.

— David Frederickson, from NYSERNet materials frederickson@nyu.edu

PSI IDS Account Service Will End in December

For the past several years, members of the NYU community have been eligible for Individual Dialup Service (IDS) accounts from PSI (Performance Systems International), permitting low-cost modem dialup to NYU-NET from locations outside New York City. However, PSI plans to discontinue its IDS service entirely this next year and NYSERNet will no longer be using PSI to provide its Internet access for New York State organizations (see article above). Therefore the IDS accounts from PSI will be terminated as of December 31, 1994.

Some individuals have reported problems dealing with PSI on business matters. If you experience such a problem, please contact Sharon Francis at the ACF (francis@acf2.nyu.edu or 998-3016), who will do her best to help you.

At this time, firm plans are not yet in place for a low-cost replacement to IDS accounts, especially for areas outside New York State (where NYSERNet does not operate). The ACF recommends that people who live outside the immediate New York City area look into getting an Internet access account from a low-cost commercial Internet provider. You can find a list of them through the NYU World-Wide Web server: from the main menu of your NYU-Internet account, select 3. Network, then 3. Lynx (or, from the command prompt of many ACF accounts, simply type lynx), then select Other Interesting Internet Resources, then Providers of Commercial Internet Access (POCIA).

— Gary Chapman chapman@nyu.edu
The Baud in Question: Save Downloading Time (and Money) with a Faster Modem

To meet the increasing demand for dial-in access from off campus, the ACF has added to its ever-expanding modem pool again, and the NYU Computer Store is ready with a selection of high-speed modems for Mac and IBM desktop computers and notebooks. We've been sorting through the dizzying array of modem technologies and standards that have appeared on the scene in the last year so that we can help you find the one that is best for you. The new standards are called V.fast, V.FC, V.32terbo, and V.34 — more about them later. Translation: Fast. Sporting data-transmission rates of up to 28.8 kilobits per second, these modems effectively double the throughput of the currently popular 14.4kBps modems, providing unprecedented data transmission speeds. In addition, these modems will send and receive faxes at 14.4kBps.

Modems that comply with the V.34 standard differ from their predecessors in that they test the quality of each telephone-line connection and are capable of slowing down in steps of 2.4kBps (previous modems stepped down in halves) to achieve the fastest possible transmission — like letting up on a car's gas pedal to reduce speed slightly. This is an important feature, since high-speed communications are very sensitive to telephone-line conditions. Some existing telephone lines, switches, and local exchanges cannot support reliable connections at 28.8kBps, but will permit them at around 24.4kBps.

A word of caution to those of you who already own 28.8kBps modems. In anticipation of the official publication of the V.34 standard in September, some manufacturers rushed modems to market based on preliminary V.34 specs. Hayes, Supra, Zoom, and Microcom all released modems based on the V.FC chip set earlier this year. Such modems are virtually identical to V.34-compliant modems, but not fully compatible. Check with the modem manufacturer whether your modem complies with the V.34 standards; if not, they may offer upgrades or replacements that are fully compliant for a nominal fee.

How Fast Is Fast Enough?

Do you really need a modem this fast? That depends. If you own a V.32-compliant modem rated at 9.6kBps or higher and use it primarily to use online services, maybe not. Most of these services offer access at a maximum of 9600Bps or 14.4kBps; and their equipment can barely handle these speeds, let alone 28.8kBps (that may, of course, change). However, if you spend a lot of time doing remote computing, or remotely controlling other computers, V.34 modems will make your life a lot easier.

Another factor to consider for Internet power users is that SLIP and PPP will benefit from using V.34 modems. Though not all telephone lines will support 28.8kBps connections, the handshaking features of the V.34 modems will be necessary to connect to these services.

Bauds, Bits, and V.bis — Some Clarifications

What do you need to know? Baud, Bps, V.32, V.32bis, V.42, V.42bis, and MNP should suffice for now. Baud and Bps (bits per second) are most commonly used to indicate the speed of a modem. Though the terms are often used interchangeably, baud is actually the measure of the number of changes of state per second on a communication channel. Several bits of information can be transmitted with each change in the channel's state. Bps refers to the number of bits that can be transferred per second. So the next time someone tells you that they have a 2400-baud modem, you can probably assume that what they really mean is that they have a 600-bit baud modem that modulates 4 bits into each change in the channel's state, resulting in transmitting 2400 bits per second, or Bps (conventionally, B stands for bit, and b for byte). So what’s the speed of this modem? Really, 2400Bps.

Modem speed, then, is easy to ascertain by looking for the Bps rating on the packages at the Computer Store. The higher the Bps rating, the faster the modem. A 9600Bps (or 9.6kBps) modem is 4 times faster than a 2400Bps modem. Suppose you subscribe to a commercial on-line service such as CompuServe, and they charge $12 per hour of connect time. A file that takes 15 minutes to download at 14,400Bps would then cost you $3. That same file downloaded with a 2400Bps modem would take a full hour and cost you $12! At a rate of only one file transfer per month, you would spend approximately $36 a year using the 14.4Bps modem, versus $144 using the 2400Bps modem. (On the other hand, many services charge more per hour for faster connections, so you’ll (continued on page 20)
CHARMmed Helices: There's More than One Way to Model a Molecule

Lately some of us here at NYU have been using QUANTA and CHARMM (the lowercase m at the end indicates the commercial version), a pair of programs recently installed by the ACF for creating molecular models. QUANTA provides the graphical user interface (GUI), which generates the models and produces the images onscreen. It is an extremely user-friendly program, readily accessible to inexperienced users; but it is also powerful enough to be useful for scientists involved in more sophisticated physicochemical investigations. CHARMM is the computational "engine," which does all the submitted calculations needed for the imaging.

Chemists have many uses for this type of software, most commonly for studying minimizations and dynamics of proteins, nucleic acids, and lipids, but CHARMM can perform a wide variety of calculations on non-biological systems as well. Examples include

- crystal calculations
- vibrational analysis
- conformational searching
- semiempirical calculations based on MOPAC
- studies of quantum and molecular mechanics
- free-energy perturbation calculations
- reaction-path mapping.

It is also useful in the analysis of a wide variety of static and dynamic molecular properties.

As an example of what the programs can do, the image on the cover was generated in the course of doing computational chemistry on the protein p21, the product of an oncogene that is normally found in all cells. An oncogene is a gene that is the blueprint for essential proteins such as growth proteins, but when mutated can cause cancer. In our research here in the Chemistry Department, we are trying to determine the mechanism by which the mutant version of p21 causes cell transformation.

QUANTA and CHARMM are available on the ACF’s Silicon Graphics workstations in the Scientific Visualization Lab (room 317, Warren Weaver Hall). Any user with an account on the SGI machines at the ACF is welcome to try the programs. Type source ~msi/quanta4.0/.setquanta and then quanta at the following prompt.

QUANTA and CHARMM make up a commercial package from Molecular Simulations, Inc., based on the software developed by Martin Karplus and coworkers at Harvard. The noncommercial software version is called CHARMM (all caps), and is not supported by Molecular Simulations.

Dr. Monaco is a research assistant professor in the Department of Chemistry (FAS).
Will multimedia change the way books are published? How can a network serve a community? What will it cost to use the Internet, as government funding is cut back? These were among the questions considered at this fall’s NYU colloquia on computers and communications.

Issues in Interactive Multimedia

A capacity audience gathered to hear Bob Stein, a founder of the Voyager Company and a leading figure in interactive media publishing, give his views on Issues in Interactive Multimedia.

Books, Mr. Stein noted, are a powerful medium in great part because they are under the control of the user. In the emergence of interactive media, he sees the return from producer-driven media — such as film and television — to user-driven media, providing, in turn, new opportunities for writers and other artists.

Mr. Stein illustrated his talk with samples from Voyager’s CD-ROM publications — including The Rite of Spring, a performance of the Stravinsky work richly annotated with discussions of the piece, its orchestration and individual instruments. The computer’s potential as a medium for expressing emotional content was demonstrated by 1 Photograph to Remember, photographer Pedro Meyer’s narrative piece on his parents’ Holocaust experiences; there is much interest among photographers in using interactive media as a new means of delivering photographic essays.

Poetry in Motion is an interactive adaptation of a 1980 film in which 24 North American poets read and discuss their work. Voyager’s first QuickTime application, the work uses multimedia to augment the film, displaying the text of each poem as the poet reads and discusses the work. With technology now permitting video to take a larger portion of the computer screen, however, images are “pushing the text off the screen,” and new-media publishers are experimenting with ways of integrating text and video. In Voyager’s treatment of Marvin Minsky’s Society of Mind, for example, a miniature Minsky perches on illustrations or moves across the “page” as he discusses the ideas underlying his book.

Mr. Stein decried the absence of critical analysis of CD-ROM publications, and of critical study of how they are actually used and “read.” He also noted that most Voyager publications at present are nonfiction, and that successful forms of fiction for new media have still to be developed.

Creating a Community Network that Scales

It’s about community; it’s not about networks. This was one of the insights shared by Kenneth Klingenstein, who is Director of Computing and Network Services at the University of Colorado at Boulder. Dr. Klingenstein is a central figure in a new project whose mission includes establishing a network for the Boulder Valley community and researching the long-term role of a community network within a national “community” of networks.

Still in its first year, the Boulder Community Network

NYU Colloquia on Computers and Communications

This popular series of colloquia on uses of computers and communications is sponsored by the ACF and the Faculty of Arts and Science, with support from Apple Computer, Inc. Individual colloquia are co-sponsored by additional university departments, depending on the topic. The colloquia are open to all NYU faculty, staff, and students. The spring series will be announced in NYU Events, the university’s biweekly calendar, and on the NYU CWIS, and flyers will be mailed to all NYU faculty. To receive an e-mail flyer, ask to be added to the ACF’s mailing list: either send e-mail to document@acfcluster.nyu.edu or call 998-3333.

Since the spring of 1993, all the colloquia have been videotaped. Copies may be borrowed from the ACF Documentation Office, Warren Weaver Hall, room 312 (998-3036).
already provides online information from city and county governments, schools, libraries, social-service groups like United Way, and newspapers and broadcast stations. It is accessible via the World-Wide Web at the address bcn.boulder.co.us. Individuals with Internet connections through their work or educational affiliations can access it free of charge, while a small subscription fee obtains access for other area residents.

An important project goal is to target groups that have so far been underrepresented on the networks. Dr. Klingenstein’s staff matches community groups that need help putting their material online with skilled individuals who provide startup support and training. The speed and enthusiasm with which use of the network is being adopted has amazed the Project staff. People from groups they had assumed to be comparatively network-illiterate — senior citizens, for example — quickly latched onto the principles and, when given the opportunity to develop their own World-Wide Web pages, became deeply involved in the task of self-definition and presentation for the Web. Community networks appear to offer important opportunities for self-realization, for democratic interchange, and for community-building.

How is the information kept up to date? Information providers sign a formal contract, and all material is posted with an expiration date, to protect readers from outdated material that hasn’t been withdrawn. Still undefined is the issue of whether a network such as this is a “public forum” and how constitutional rights of free speech might obtain.

Cost savings are among the benefits to agencies providing information to the community network. For instance, the United Way no longer has to print its Red Book of information on services provided by its constituent organizations, and with online publication, the material can readily be kept current; word-search capabilities also make the information easy to find and use.

**Pricing Internet Access and Use**

With the National Science Foundation (NSF) withdrawing its funding of the Internet backbone, and scaling back its subsidies for the regional networks, we’re likely to see major changes in the industrial organization of the Internet as it becomes increasingly privatized. How will Internet access and usage be priced in the coming years?

“Usage-based pricing,” predicts Hal R. Varian, the speaker at this fall’s third colloquium and a Professor of Economics and Finance at the University of Michigan. Will costs to individual users and to universities rise significantly? In the near term, probably not. Professor Varian reports that the NSF now funds the Internet backbone to the tune of about $12 million a year, and subsidizes regional networks with another $8 million. Since roughly 12 million people nationwide now use the Internet, per-person costs aren’t much.

In the coming years, however, with Internet traffic increasing 6 or 8 percent a month — reflecting both new users and the growing popularity of new multimedia applications like teleconferencing and Mosaic — the demand for bandwidth will rise dramatically and will push the need for some sort of usage-based pricing. While a text-based Internet transaction like sending e-mail or transferring an ASCII file has an insignificant impact on traffic, the transfer of sound, images, and video is quite a different matter. For example, Professor Varian estimates that a 1-megabyte file can hold the text of a 700-page book, but only a few images, or a three-second video.

Professor Varian and his associates are studying a “smart-markets” model of Internet pricing, in which users — or, actually, their computers — “bid” for Internet access and request a priority-level for each transaction. For example, you might request a high priority for real-time video, but tolerate a small delay for e-mail. Under this model, the price paid for an Internet transaction would depend on how much load it is putting on the system. A high-priority item would be transmitted at a higher rate when traffic is heavy, and less when traffic is light. The price of transferring text would continue to be negligible.

Professor Varian’s recent research has focused on the economics of information networks. Several of his papers and related works are accessible on the Internet. Using Lynx or Mosaic, connect to http://gopher.econ.lsa.umich.edu/ and see the items under Economics and the Internet.

— Estelle Hochberg and David Frederickson hochberg@nyu.edu • frederickson@nyu.edu

**ACF HelpLine Q&A**

**Q:** Why don’t my ACF username and password work when I try to log on to BobCat?

I never had to supply this information before.

**A:** Try using “bobcat” as your username. The new version of BobCat, unlike the old, needs to have you log in — but it doesn’t need your real username. You probably missed the onscreen instructions that tell you to enter bobcat (all lowercase) at the login prompt. This should take you directly into the new and improved BobCat system. If you enter anything other than bobcat, you’ll be prompted for a password and won’t get anywhere.

— L. Barnett

**Call the ACF HelpLine at 998-3333**
Geocoding at the ACF's Statistics and Social Science Group

Frank LoPresti
lopresti@nyu.edu

Within the academic research community, geographic information systems (GIS) are proving useful in more and more fields. GIS is a logical next step in data management and analysis. Thirty years ago, researchers were satisfied to use a Fortran or Cobol program to manage and explore data sets — a laborious, technical undertaking. Then came PCs, along with better editors and statistical packages. At the same time, Lotus and dBase-type software gave us powerful tools to organize larger, more complex data sets. But all of these methods yielded only tables and graphs — not always the most revealing displays.

A great deal of data can be usefully and revealingly located on a map. Analyzing relationships of data at one point on a map to similar data or to other features found on a map was attempted early on. The development of these analysis tools has led to the modern GIS software now being used in science, public administration, marketing, forest inventories, flood and fire modeling, anthropology — a broad and growing range of fields. NYU's Academic Computing Facility now supports the most popular GIS packages, including MapInfo, Atlas GIS, and ArcInfo.

GIS works with map features — points, lines, and polygons — and uses the associated coordinates to position these objects on a map. These features can represent mailboxes, houses, streets, lakes, etc. From a different source, we might have data we want to attach to some of the objects making up our map. These would be in files (perhaps in a database format), such as census summary data with counts at various levels, for instance, counts of persons by income level within census tracts. Or perhaps the attribute data are measurements of water flow at many points in rivers and streams at various times of the year. In the first case, we might be interested in intelligently locating future services using population density. In the second case we could be studying flood plains and land use.

Geocoding

Before the researcher can do any GIS work — spatial analysis or creating thematic maps — the attribute data have to be linked to map features. This linking is known as geocoding. In the past, it took quite a bit of tedious work.

The ACF's Data Base Archive (DBA) acquires and stores data files for instructional and research purposes at NYU. Assistance in the use of these data files is provided by ACF staff to NYU faculty, researchers, and graduate students. The DBA currently holds and catalogs some 700 studies represented by over 2,400 data files. More are being acquired continually at the request of researchers at NYU. For further information on the DBA’s services, or for help in making use of them, please contact ACF consultant Frank LoPresti (998-3398) or Bob Yaffee (998-3402). Also see “Gopher Access to the ICPSR Catalog” in the January 1994 issue of this newsletter. Full descriptions of DBA holdings may be seen by typing HELPME DATABASE on WYLBUR, or help database on the ACFcluster, or by speaking to the ACF's Statistics and Social Science staff.

Frank LoPresti heads the ACF’s Statistics and Social Sciences group in Tisch Hall.
The screen shots at right show windows from two modules from Wessex, designed to bring data seamlessly into the GIS program Map Info. The Loading Program (top) is used to process the Census Bureau’s TIGER files for geocoding, as described in the accompanying story. Wessex Pro Filer (bottom) extracts subsets of data from the Census Summary Tape Files (STF) and produces tables as ASCII files or in dBase format. The researcher can specify which geographical areas (state, county, congressional district, etc.) to use, and which census variables and summary levels to include in the table. Usually one would first create a map file with the Loading Program, then select the data in Pro Filer. There, each of the numerous Available Menus can bring up several variables (middle column); the right-hand column lists the variables selected. In Map Info, the researcher can then edit the map and join the data to it; since Wessex has made sure the key values are compatible in map and data, the joining is essentially seamless.

An article by Thrall and Thrall in the May 1994 issue of Geo Info Systems details three methods of geocoding: address-matching, point-creation, and joining and boundary-matching.

Address-matching is the process of linking attribute data to points along streets. In GIS, streets are lines — perhaps straight, perhaps curved. The map has the lat/lon (latitude and longitude) details needed to position the street, and also the range of addresses on the segment. In address-matching, the range of addresses is proportionately distributed along the segment so we may determine the approximate location of a given address. Address-matching software also attempts to put the “house” on the right side of the street, using the range information attached to the segment in question. Intelligent geocoding software also tries to resolve the disparity between a target address spelled one way and a different spelling on the map.

Geocoding by point-creation simply matches data in a target list with points on a map that has lat/lon information. Changing the lat/lon coordinates to the internal GIS map coordinates is the main function of point-creation geocoding. Additionally, when no perfect match occurs, it may reposition (snap) some of the points to the nearest feature. An example of this is in building a tax map using Global Positioning System hardware coupled with a PC to locate houses and inventory structures along roads on map layers. Some cartographers do not consider joining and boundary-matching to be geocoding at all, as they needn’t involve addresses. Joining is the standard procedure — using keys (e.g., numbers or letters) to join data on a target table (some table of attribute values) and the objects on the map. The target side of the join must have no duplicate keys. For example, the map layer delineating census tracts would have a keyed field for tract number. The target table could have the tract number and attribute fields for counts of, say, children. Joining these would allow the researcher to produce thematic maps displaying the density of children, where dark shading within a tract would represent a high count of children per square mile and light shading would show a lower density. Boundary-matching uses boundary name as key to match. The geographic file will be analyzed to find a centroid location, a spot to which the data are appended.

Wessex Pro Filer — Linking Data to Map Features

The ACF has recently acquired Wessex Pro Filer, a combination data and software product that allows researchers to access demographics from the 1990 U.S. Census with ease. The Pro Filer CD-ROM disks include the data from the Census Bureau’s STF-1A and 3A files (from the short and long questionnaires, respectively) for the entire nation. Using menus, the researcher extracts particular fields and limits the geographic region and summary level. The field names

Academic Computing and Networking at NYU November 1994 17
are presented in understandable words, such as age or income — not as unintelligible parts of matrices such as the “P004 matrix, variable 028” in Census documents.

These extracts are saved in dBase format with the geographic links that allow seamless geocoding, vastly speeding up a formerly tedious task. These data and extraction programs are now available to NYU researchers at the ACF’s Social Science group in Tisch Hall.

**Doctor Dolittle**

The GIS-related software Doctor Dolittle (yes, only one o), is also now available at the Social Science group. This software “tames” TIGER files, making them useful for Atlas GIS. A second module allows conversion to the MapInfo GIS format.

TIGER (which stands for the Census Bureau’s Topologically Integrated Geographic Encoding and Referencing data set) is part of a cooperative effort between the Census Bureau, the U.S Coast Guard, federal, state, and local agencies, and the private sector to build a map base for the entire United States. The TIGER data delineate such attributes as roads, schools, hospitals, power lines, railroads, rivers, parks, street addresses, etc. The TIGER Line File is a series of line segments with lat/lon coordinates for beginning and end points, shapes, and various descriptive information of these attributes. With it comes the potential to form boundaries for all Census-related geography for 1990 and for historic information from 1980.

Doctor Dolittle will link the TIGER line segments to form polygons for census geographic levels for 1980 and for 1990. The problem is that TIGER CD-ROMs just don’t give the topology we might expect. Line segments aren’t linked to give polygons for summary levels; thus the map we get will look correct, but the GIS will not “know” we are inside a census tract, since it is not a closed polygon. Using Doctor Dolittle, we could ask for counties from the 1990 census; the line segments needed to create the tract polygons for 1990 would be selected and closed, and the tract names would be applied as labels. A file with those newly created polygons would then be available for use with Atlas or GIS for Map Info.

With these new tools available at NYU, GIS takes a big step away from the domain of high-overhead, exotic software. Computerized mapping is well on its way to becoming one of the standard modern research methods.

**LISREL for Windows Raises Statistics Stake at the ACF**

For more than a decade, researchers and statisticians have come to rely on successive versions of LISREL (short for Linear Structural Equation Relations) for high-end statistical modeling. The authors of this package, Karl Joreskog and Dag Sorbom, have now produced version 8 for Windows, and the Academic Computing Facility has acquired LISREL8W for use on its PCs.

LISREL8W performs confirmatory factor analyses, structural regression analyses, regressions between latent variables, path analyses, path analysis between latent constructs (factors), and comparisons of covariance structures across samples, among other things.

Several new features distinguish this version of the package from earlier ones. For one thing, the Simplis programming language is incorporated within the package, so that the commands can be issued in ordinary English. Simplis allows one to input the problem title, the observed variable list, the covariance matrix, the sample size, and the structural equation, whereupon the LISREL output will be generated. These files have .spl suffixes and are identified as such by LISREL8W.

New path diagrams can be generated with LISREL8W. When you insert the Path Diagram command, a path diagram is produced, depicting the model being estimated. The path diagram file may be saved as a Windows bitmap (.bmp) file and imported into your favorite word processing software.

Unlike previous versions of LISREL, the new output contains reliabilities of the measurement equations and can model correlations between the exogenous and endogenous errors. It calculates the reliabilities of the factor loadings by squaring them, and gives them along with the output from the measurement model.

Criticism had abounded in the user community that earlier versions of LISREL had only a handful of goodness-of-fitting criteria; many people had difficulty getting their models to fit with those functions. They will be pleased that the new version has added many more. Among the new Goodness of Fit functions are the population-discrepancy function (pdf), which assesses the error of approximation in the population, and Steiger’s Root Mean Square Error of Approximation, which measures inflation per degree of freedom that accompanies that pdf. A number of information measures — such as the Akaike Information Criterion (AIC), Corrected AIC, and Expected Cross-Validation index — are also included. Other Fit indices include the Normed Fit index, the Non-

(continued on page 20)
Periodical Abstracts Goes Online: Now Available from Home or Office PC

CD-ROM indexes to periodicals and newspapers are among the most popular reference sources in Bobst Library. Located in the Library's three Reference Centers (1st, 6th, and 9th floors), these networked CD-ROMs are frequently all in use during peak library hours. Now, researchers with ACF network accounts can get a head start on their research from their home or office by using the online version of Periodical Abstracts, a general periodical database providing journal citations (including short abstracts) covering virtually all disciplines.

Periodical Abstracts provides access to citations from over 1,650 periodicals and transcripts of television news broadcasts from 1986 to the present. The citations are displayed in reverse chronological order (most recent first). Search commands are simple, and search results can be sent to your e-mail address or downloaded.

Periodical Abstracts is just one of the EUREKA databases. Other choices include HAPI (Hispanic American Periodicals Index), Handbook of Latin American Studies, Avery Index to Architectural Periodicals, Anthropological Literature, and others. To gain remote access to all of the EUREKA and other Library resources from your ACF account, follow the instructions in the box at right.

Changes in Bobst Library's DIALOG Access: More Databases, More Powerful Searching

Bobst Library continues this year to offer access to DIALOG databases. This summer, the Library expanded its DIALOG database access to include many more resources. Users can now search over 300 databases, covering virtually all subject areas of interest. Besides adding new sources, the DIALOG Information System also became more powerful. Now DIALOG users can conduct searches of related databases simultaneously, choose from different record formats for downloading, and target search results according to relevance or importance.

DIALOG provides access to four major types of databases: citation (usually with abstracts provided), full text, and others. To gain remote access to all of the EUREKA and other Library resources from your ACF account, follow the instructions in the box at right.

Connecting to the RLIN/EUREKA Databases

To connect to Bobst's RLIN and EUREKA databases, follow these steps:

1. From an ACF NYU-Internet account: at the main menu,
   Choose Extras
   From an EMIS account on the ACF-cluster: at the main menu,
   1. Choose 5 Software
   2. Choose 2 Communications
   3. Choose 3 RLIN
   4. Choose 1 RLIN
   The system automatically logs you on and places you within the BIB database.
   To select any of the other databases, simply enter the command Choose followed by the three-letter database code, e.g., Choose PRA (to search Periodical Abstracts).
   Guides for searching EUREKA are available at Bobst Reference Desks (1st, 6th, and 9th floors).

New Databases and Electronic Resources at Bobst Library, Fall 1994

New databases and electronic resources are up and running in the reference centers (1st, 6th, and 9th floors), the Avery Fisher Center (AFC; 2nd floor) and the Electronic Resources Center (ERC; B level) in Bobst Library. Below is a list of the new resources. For a complete list of electronic databases and resources in Bobst, consult the Bobst Library portion of the NYU CWIS.

Contemporary Authors (Ref 1)
EconLIT (Ref 6)
Encyclopedia Britannica (ERC)
Ethnic Newswatch (Ref 1, 6)
EuroCat (Ref 6)
FBIS (Ref 6)
Le Robert Electronique (ERC)
Muse (Ref 1, AFC)
Music Index (Ref 1, AFC)
Oxford English Dictionary (ERC)
SocioFile (Ref 6)
Statistical MasterFile (Ref 6, 9)
UN Index (Ref 6)
Wall Street Journal, full text (Ref 1,6)
The Baud in Question (continued from page 12)

have to figure out what the best deal is.) For frequent commercial on-liners, a faster modem can quickly offset its higher price tag.

Speed is not the only consideration when buying a modem, however. You should see on the package that the modem supports several standards — probably V.32, V.32bis, V.42, and V.42bis. V.32 is the CCITT (now called IUT-T) standard that generally indicates 9600Bps-capable modems. V.32bis refers to the standard for modems operating at 14400Bps, or 14.4kBps. V.42 is the standard for error control and correction (including standards MNP-2 through MNP-5 error-correction), which means the modem can achieve a higher accuracy rate of data transmission, which is important when it is connected to inferior or noisy telephone lines. If your modem lacks these error-control features, a noisy line can wreak havoc on your communications — a frustration easily minimized by knowing your Vs.

MNP-5 and V.42bis indicate data-compression standards, which means that the modem can squeeze more data through the communication pipeline, thereby achieving a higher transmission or Bps rate. A modem rated at 14,400Bps with V.42bis compression is theoretically capable of throughput of up to 57,600Bps when connected to another modem meeting the same standards. In general, the more Vs you see the better, as these modems are more powerful, flexible, and compatible with other modems. Seasoned telecommunicators usually prefer these "Swiss Army Knife" modems.

Be wary of cheap, off-brand modems, as they are often missing some Vs. We strongly recommend that you stick with the major manufacturers. The money you save with an off-brand modem can be quickly offset by time-consuming and frustrating interrupted file transfers, retransmission of data, and incompatibility with other modems.

Confused? We're not. Stop by the NYU Computer Store and ask for a 28.8kBps modem, and rest assured that it will be V.34-compliant and surprisingly affordable. Expect to pay around $370 to enter the fast lane. Or, if you prefer a more leisurely cruise down the information highway, we have fax-modems capable of 2400, 9600, 14,400 and 19,200Bps, loaded with Vs to suit your needs.

— Bryan E. Leonard
leonardb@is.nyu.edu
Bryan E. Leonard is Administrative Manager at the NYU Computer Store

LISREL (continued from page 18)

Normed Fit index, the Parsimony Normed Fit index, the Parsimony Goodness of Fit index, the Comparative Fit index, Incremental Fit index, and the Relative Fit index. A critical N statistic is also added to show the sample size that would make the obtained chi-square just significant. With this version of LISREL there are many different fitting criteria by which to assess your statistical model.

The new version of PRELIS 2.1 that accompanies LISREL8W can handle non-normal data. It can process rank correlations, such as the Tau-C, by generating empirical non-normal standard errors with its new bootstrapping option. Moreover, Scientific Software maintains that another new option in Prelis may be used to impute means to missing data.

Running LISREL8W is very easy. You may edit your program as a Simplis program with the simple wordprocessor in Windows and save it in the LISREL8W subdirectory with the suffix .spl. Then click on the button that says Run LISREL and the program is executed. If you have placed the statement path diagram at the end of the program, a path diagram will automatically appear, indicating that the program has run. If you would like assistance with the theory or programming of LISREL, call me at 998-3402 for an appointment.

Altogether, many old deficiencies are remedied and new power is added to a user-friendly but sophisticated statistical package that permits analysis of covariance structures on the popular Windows computer platform here at the ACF.

— Robert A. Yaffee
yaffee@nyu.edu
Dr. Yaffee is a member of the ACF’s Statistics and Social Sciences group.
Important ACF Telephone Numbers
ACF HelpLine  998-3333  
Account Information  998-3035  
Computer Documentation  998-3036  
ACF Innovation Center  998-3044  
Statistical Consultants  998-3434  
Computer Labs:
  14 Washington Place  998-3457  
  Tisch Hall  998-3409  
  Education Building  998-3421  
  Warren Weaver Hall  998-3456  
  Third Ave. North Res. Hall  998-3504  

Dial-in Access to ACF Computers
(Via NYU-NET, NYU’s campus-wide network.)
If calling from Dial  For (bps)  
  Off Campus  995-3600  300-2400  
  995-4343  9600, 14400  
  995-4335*  300-1200  
*This number is recommended if you are using an older modem that has no error-correction.

Exceptions to regular hours: confirmed holiday schedules will be posted via our online news and bulletin board facilities, and ACF offices in Warren Weaver Hall are closed on University holidays.

NEW YORK UNIVERSITY
Washington Square Center

Guide to ACF user work areas and other facilities

1. Warren Weaver Hall, 251 Mercer St., 3rd floor
2. Tisch Hall, 40 W. 4th St., lower concourse (Rooms LC-7 and LC-8)
3. 14 Washington Pl., basement
4. Education Building, 35 W. 4th St., second floor
5. 715 Broadway (IBM tapes only)
6. Third Ave. North Res. Hall, 75 Third Ave., level C3

ACF microcomputer labs are at (2), (3), (4), and (6). NYU Trolley and NYU Shuttle routes include (6): Mon.–Thurs., every 15 minutes (8:00 am–11:00 pm); Fri. and Sun., every 30 minutes (6:00 pm–1:00 am) during the academic year.

Regular Fall Hours at ACF Sites

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>14 Washington Place*</td>
<td>closed</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 5:30 pm</td>
</tr>
<tr>
<td>Tisch Hall*</td>
<td>closed</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 5:30 pm</td>
</tr>
<tr>
<td>Education Building*</td>
<td>closed</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 11:30 pm</td>
<td>8:30 am – 5:30 pm</td>
</tr>
<tr>
<td>Third Ave. North*</td>
<td>10:30 am – 1:30 am</td>
<td>10:30 am – 1:30 am</td>
<td>10:30 am – 5:30 pm</td>
<td>10:30 am – 5:30 pm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consultant Hours:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>14 Washington Place</td>
<td>closed</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 5:00 pm</td>
</tr>
<tr>
<td>Tisch Hall</td>
<td>closed</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 5:00 pm</td>
</tr>
<tr>
<td>Education Building</td>
<td>closed</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 9:00 pm</td>
<td>9:00 am – 5:00 pm</td>
</tr>
<tr>
<td>Third Ave. North</td>
<td>10:30 am – 10:00 pm</td>
<td>10:30 am – 10:00 pm</td>
<td>10:30 am – 5:30 pm</td>
<td>10:30 am – 5:30 pm</td>
</tr>
</tbody>
</table>

*Available to general users from 8:30 am to 1:00 pm, Mon. through Fri., and to priority access account holders during all hours of operation.

**Available to general users and priority access account holders during all hours of operation.