Transforming Research Through Technology

By Pauline Stieff

Teaching, learning and research has greatly benefited from the use of technology, particularly in the last decade. Yet, the struggle to gather, compile, manage and analyze large amounts of data still remains a challenge in certain disciplines. Also, the expansion of the Internet and more sophisticated technology has created the cumulative problem of how to capture and utilize increasing amounts of complex and related information.

Informatics is the processing and analysis of large amounts of information for storage and retrieval. Supercomputers are the powerful and complex machines used to manage informatics. All disciplines can benefit from the use of informatics since it removes many of the logistic and computational burdens traditionally associated with conducting research or collaborating with colleagues. The limitations of using laboratory space and special instruments for collecting data as well as and the corresponding delays and dilemmas of compiling and analyzing research results are eliminated or greatly reduced. Researchers in different locations can send and receive data and communicate via videoconferencing in addition to the traditional electronic communication. In the past, such interactions were limited since the transmissions of large amounts of data delayed communications.\(^1\)

Bioinformatics is the field of science in which biology, computer science and information technology merge to form a single discipline. The ultimate goal of the field is to enable the discovery of new biological insights as well as to create a global perspective from which unifying principles in biology can be discerned. Bioinformatics has expanded to include many different sub-disciplines including: evolutionary biology, protein modeling and genome mapping. It has “transformed biology from a lab-based science to an information science as well.” As an emerging field, bioinformatics promises to lead to advances in understanding basic biological processes and, in turn, advances in the diagnosis, prevention, treatment and prevention of many genetic diseases. The implications of such an effect in both science and health care are immeasurable.\(^2\)

Research in bioinformatics as a part of informatics can serve as a model for computation in other disciplines. Tufts as a multidisciplinary university can uniquely benefit from the use of such technology. A number of our faculty (a few are profiled in the following article) are involved in collaborative projects involving supercomputing and informatics-related technology. Many federal agencies, including the National Institutes of Health and NASA offer information as well as grant programs on the use of informatics. The Tufts University Research Network (TURN) is promoting interdisciplinary research and plans to submit proposals for such programs this year.

For more information on Bioinformatics, go to: www.nih.gov/about/director/060399.htm or www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html

\(^1\) This summary of information was extracted from a paper presented by Dr. Robert Chapman of the Tufts Dental School at the Academic Affairs Committee Meeting on January 20, 2004

\(^2\) From “A Science Primer,” National Center for Biotechnology Information, November 2001

Pauline Stieff, Grants Specialist in Academic Technology, is a writer with extensive development experience in the non-profit sector.
Innovative Roles of Technology In Research, Part One
By Pauline Stieff

Bruce Boghosian is a professor in both the Mathematics and Computer Services Departments on the Medford campus. He was also a Research Associate Professor at Boston University Center for Computational Science and Physics for almost 10 years. After earning a Master’s degree in nuclear engineering from MIT and then a PhD in Applied Science and Engineering from the University at California, Davis, Boghosian worked as research scientist at a Cambridge start-up company called Thinking Machines Corporation. This company pioneered parallel computation for scientific applications.

Boghosian came to Tufts in 2000 and was attracted by the university’s unique research environment. “Some schools offer faculty the opportunity to develop close relationships with undergraduate students, he says. “Other schools offer faculty the opportunity to pursue bold research agendas and be entrepreneurial. Tufts is one of the very few schools that can do both well,”

According to Boghosian, over the past fifteen years, the advent of parallel computing has allowed users to distribute computational tasks over multiprocessors or local clusters of workstations. Grid Computing aims to go one step further and allow users to spread their tasks over many multiprocessors and a wide geographic distribution. The most famous example of this is the SETI (Search for ExtraTerrestrial Intelligence) project at Cornell, which offers a special screen saver to the public. When a user’s computer is cycling idly, the screen saver kicks in, downloads some of SETI’s data that needs to be analyzed, and starts crunching on it. With tens of thousands of users throughout the world, SETI is able to harvest enormous amounts of computer power that would otherwise have been wasted. This is sometimes called “cycle harvesting.”

More realistically, not all computational tasks can be partitioned as finely as SETI’s, so there needs to be some controllable combination of distribution and locality. Boghosian continues, “for example, it would be inefficient for our lattice simulations of fluid dynamics to be spread all over a geographically distributed grid because the lattice computation is so tightly coupled — neighboring lattice points need to communicate frequently. On the other hand, we could easily imagine setting up the initial conditions on one machine, evolving the fluid in time on another machine, periodically post processing output on another machine, and rendering graphical output on yet another.”

Boghosian is also a Visiting Fellow at the Centre for Computational Science of University College London, where he participates in a project called RealityGrid, aimed at applying grid computing to problems in materials science and fluid dynamics. This project is funded by the British Engineering and Physical Sciences Research Council that mandated it should involve both computer scientists and application programmers. Therefore, the development of Grid Computing in the UK is heavily influenced by the needs of application programmers. The US program is more tilted toward computer scientists and computer manufacturers.

As part of the RealityGrid, Boghosian’s group has a long-term project aimed at studying the formation, evolution and decay of vortex knots in fluids. A familiar example of a vortex core is a tornado; the vortex core is more or less vertical, and the air swirls around it. A more interesting example is a smoke ring, since that has a circular vortex core. Such vortex cores can tangle together, and viscosity can cause them to interconnect into more complicated knots and links. In fully developed turbulence, vortex knots and links continuously appear, evolve for a certain lifetime, and decay away, to create the turbulent steady state. Boghosian and his team use a combination of theoretical analysis and high-performance simulation to study the dynamics of individual vortex knots and links.

Boghosian along with Peter Coveney of the Department of Chemistry at University College London, and Stephen Pickles of the Department of Computer Science at the University of Manchester cocreated a demonstration at Supercomputing 2003 in Arizona last November in which they simulated a phase-separating amphiphilic fluid — a mixture of oil, water, and surfactant, separating due to the surface tension of the oil/water interface, but ultimately stabilized due to the surfactant. Such mixtures can form fascinating “liquid crystals,” including a very elusive
“gyroid” phase that has been seen in the laboratory, but has been notoriously difficult to simulate. They were able to evolve a homogenized ternary mixture of oil, water and surfactant to a stable gyroid phase, and study the evolution of dislocations — imperfections in the liquid crystalline structure. Such a simulation must necessarily be done on a large grid, since small grid sizes will suppress the formation of the dislocations.

Boghosian’s team ran simulations on grids of various sizes on a variety of machines, including the Pittsburgh Supercomputer Center, and NCSA. They generated graphical output from these simulations on machines in London and Manchester. As the output was rendered, Lucas Finn, a graduate student in Mathematics, was able to alter parameters in response to what appeared on the screen, modifying the simulation “on the fly” — a technique called “computational steering.” All of this was broadcast over the AccessGrid to the Supercomputing meeting in Arizona, and numerous other sites throughout the world. The presentation was awarded the HPC Challenge Award for “Most Innovative Data-Intensive Application.”

To be scalable at a scientific computing performance level, many existing codes will need to be re-engineered, so that they may orchestrate the execution of numerous sub-components, over a wide geographic distribution. Like the advent of parallel computing ten years ago, or of vector computing twenty years ago, this will require some adjustment in application programmers’ thinking, but the results will be well worth the effort expended. Beyond traditional scientific computing and mathematical modeling, grids will become useful tools for facilitating database access, and for automating experiments.

By the summer of 2004, Boghosian hopes to have componentized all of the codes that he and his students use, so that we can begin to use them more effectively, treating the resources available to them as a Grid. Of course, says Boghosian, the goal of such computing is to learn new science and new mathematics, and we are looking forward to the results of numerous numerical experiments on vortex knots and turbulent fluids.

Tufts researchers have extensive expertise in high-performance and parallel computing, in many different departments. AT&TCCS are offering parallel computing as a campus resource, and so Boghosian’s team may soon get their own AccessGrid node. He believes that Tufts could play an important role in the NSF’s TeraGrid project, which later this year, will encompass 20 teraflops (trillions of floating-point operations per second) of computer servers, distributed across five sites.

BoGHOsian bElieves, within the next few years, all leading edge scientific computation will be done on arrays of thousands of processors, allowing the execution of “petaflops” — quadrillions of floating-point operations per second.

Dr. Laurie Baise is an Assistant Professor in the Department Civil and Environmental Engineering (CEE). After receiving a BSE from Princeton University, she went on to receive two Master’s Degrees from the University of California at Berkeley, one in Geotechnical Engineering and the other in Geology and Geophysics. She also received her PhD from Berkeley in Geotechnical Engineering. Dr. Baise’s graduate work focused on the study of earthquakes and she received several fellowships and awards including those from the National Science Foundation and the California Universities for Research in Earthquakes Travel Award to attend the 12th World Conference in Earthquake Engineering in New Zealand.

Baise came to Tufts in 2001 and was impressed by its size, excellent students, and the congenial environment as well as the close connection between the schools of liberal arts and sciences and engineering. Tufts provides undergraduate students with many research opportunities both within GIS (Geographical Information Systems) and other Engineering disciplines. President Bacow’s Summer Scholars Program also provides more venues for undergraduates to participate in discipline-related research. Like Dr. Boghosian, Baise believes Tufts provide a unique research environment for teaching and learning.
She is the recipient of both AT’s 2003 APT Grant for a Web Interface for Two and Three Dimensional Visualization of Subsurface Data and the 2003 Berger Family Technology Transfer Endowment for the Digital Boston Geotechnical Database for Research, Teaching, and Technology Transfer in Education along with colleagues from Geology the Tisch Library and AT.

Through collaboration, both grants will develop a digital online database of geotechnical data in the Boston area for use in research and education at Tufts University. The project focus will be on the creation of a digital library resource (as a part of Tufts Digital Library) that can be navigated using innovative technology for two-dimensional (2D) and three-dimensional (3D) visualization. This project will be accomplished by developing a GIS and web interface for accessing geotechnical data and geologic and historic maps. Project staff will develop GIS in parallel with a set of teaching modules for undergraduate and graduate courses at Tufts and will build on previous Tufts initiatives including 1) The Digital Library Collections’ Boston Streets Project, 2) the compilation of geotechnical data led by Professor Lewis Edgers and published in the Boston Society of Civil Engineers Journal, 3) an on-going research project which is developing a database of geotechnical data for Boston to assess seismic hazard, and 4) on-going research mapping stratigraphy and sea level history in marsh deposits around Boston.

The teaching modules will enable students and researchers to highlight the advantage of digital data and visualization through GIS in engineering, education and research using the data retrieval/visualization tool (GIS for geotechnical data). Each module will direct the student to use the GIS tool to answer specific questions about the subsurface conditions in Boston and how those conditions affected the development and resulting infrastructure of Boston. The GIS for geotechnical data will provide a powerful educational tool for teaching important engineering skills to students within the context of geology and Boston history. In collaboration with Tisch Library staff, instructional modules will be developed on accessing and using the digital data for the Tufts community and for other in the Boston area.

Baise’s primary goal is to link this extensive dataset of geotechnical data with a tool that will make it easier to use in the classroom. Her work could be expanded to include other disciplines in the future. Baise states, “Visualization is a very powerful teaching tool, especially when your data is underground. Since the data used is three-dimensional, this results in a few challenges for GIS since ArcView (the software that we are using primarily) is really two-dimensional software. One of the issues is that our database has one-to-many relationships which are difficult to handle in ArcView. Other projects may be able to benefit from how we handle these challenges. In terms of 2D and 3D visualization, geology can clearly benefit from these tools. In terms of the dataset to which we are linking with historic maps, the entire project has historical relevance and therefore could be interesting to history students as well as to urban and environmental planners.”

Pauline Stieff, Grants Specialist in Academic Technology, is a writer with extensive development experience in the non-profit sector.

For more information about resources discussed in this article, go to:

RealityGrid
• www.realitygrid.org

SETI:
• wwwSETI.gov

PSC:
• www.psc.edu

NCSA:
• www.ncsa.edu
• www.ucl.edu

GIS websites:
• www.state.ma.us/mgis
• www.bigdig.com
• www.mapjunction.com/index.htm

This two-part series will continue in the next issue featuring additional faculty using technology in research.

May, 2004

In the Next Issue of Innovations in Learning Newsletter:
Legal and Ethical Issues Using Technology
In research, often the most effective collaboration of talent and resources unites contributors in a partnership across great distances. One institution provides the intellectual resources, perhaps another offers access to equipment, and another contributes experimental subjects. These partners are geographically removed so that frequent meetings, let alone face-to-face collaborative investigations, are not options. In circumstances like these, electronic cooperative investigation in a shared virtual space is often one option for supporting the research team. Until recently, these virtual “collaboratories” have not been feasible for institutions without an extensive technological infrastructure. Advances in digital imagery, communications and computing hardware and software, collaboratories have become increasingly popular. What remains to be seen is how effective they are. What can be done to minimize the weaknesses and capitalize on the strengths of these virtual partnerships?

Three researchers from the University of North Carolina – Chapel Hill set out to discover if scientific research that is distributed across distances can produce high quality results, if the capabilities afforded by collaboratories outweigh their disadvantages, and how this new method changes the scientific process. They set up a simple test using a single-user nanoManipulator (nM), a device that provides visualization interfaces to a local atomic force microscope. Using the nM, scientists can interact directly with samples ranging in size from DNA to single cells. Local computers support the collaborative manipulation and exploration of data. The scientists can dynamically switch between working together in a shared mode to working independently in private mode. Collaborators can perform almost all operations synchronously. They are connected via video conferencing, desktop application sharing via Microsoft NetMeeting (available on standard Windows operating systems), wireless headsets, and speakerphones.

The experiment involved twenty pairs of students divided into two groups to create lab reports on scientific research. Each group spent time working both face-to-face and remotely. In the course of their work, the groups operated equipment, recorded data, performed data analysis, drew conclusions, created hypotheses, and prepared a formal report. The students generated a total of 40 lab reports that were graded blindly — the graders were kept unaware of whether the report had been generated by a group working face-to-face or by one collaborating remotely. Participants were also interviewed individually to gain further perspective on their experience.

Average lab scores were identical, suggesting that working remotely did not impair lab performance. The data even indicated that it might have been an advantage to collaborate remotely prior to working face-to-face. Those that completed the work in this order scored significantly higher on the second task than did those who collaborated face-to-face first. The researchers used their interview data to provide a more detailed explanation of how this happened. They discovered that, although the remote collaboration posed obstacles to the ordinary methods of collaboration, the solutions that students improvised were adequate, even advantageous, when compared to the face-to-face work. Because the working relationship was less personal when it took place remotely, participants reported that they tended to be more focused on the task at hand and less distracted by social interactions. Because they received fewer social clues about their partner’s focus and intentions, they were prompted to communicate more actively in order to keep in touch during the investigative process. Working remotely afforded both individuals an opportunity to experiment independently as well as collaboratively. The researchers quote one study participant:

Sometimes when you’re working side by side with somebody, you have to deal with ‘Well, you’re looking at [the data] from a different angle than I am, and so you’re seeing a different perspective there.’ Now [working remotely] we could both of us be straight on, having the exact same perspective from where we’re sitting. It made it easier, (pg. 5)

The response, however, was not wholly positive, however. The flip side of not being able to read each others’ social clues during the collaboration was that often the context for communication was lost; it wasn’t easy to infer why someone was asking a particular question or what they meant by a given observation. The increased need for interaction was not always seen as a benefit. To some, it hindered the progress of the research objectives.

What might we learn from this? The researchers state that the experiments performed by the two groups were not identical, so it cannot be concluded with absolute certainty that the advantage attributed to working remotely before working face-to-face will always be observed; it is possible that aspects of the tasks inhibited collaboration. The study also purposefully omitted aspects of the scientific process, such as problem formulation, research design, and research dissemination. In future studies, the authors intend to investigate whether professional scientists report similar results. In the meantime, collaboratories continue to address the pressing needs of partners who share research interests, if not research facilities.

**by Daniel Cogan-Drew**

*Daniel Cogan-Drew is the Technology Coordinator for the Education Department and a former high school and post-secondary teacher.*
What is an ePortfolio?
ePortfolios are web-based or electronic portfolios. Although static, linked web page portfolios have been on the Internet for many years, ePortfolio software applications are dynamic, database-driven environments that will be connected to Course Management Systems (such as Blackboard), Student Information Systems, and available indefinitely to graduates. These systems will enable students to customize different views of the portfolio for different audiences, invite feedback from faculty and peers, post self-reflection on portfolio contents, and capture the revision process for work-in-progress. Although still in the early stages of development, ePortfolios are conceptualized to have many uses in higher education:

• Students can showcase their learning in a single course or throughout a program of study, as well as ‘take’ their portfolios with them upon changing schools or when graduating.
• Graduates can demonstrate how they have met program or certification requirements, as well as present skills and accomplishments for employment or career advancement.
• Faculty can collect and organize student work and course materials as well as research data and reports and other professional information for tenure or peer-review.
• Academic advisors can review student coursework, faculty feedback, and help students connect university and program learning outcomes to specific course activities.
• Departments, programs, and schools can access faculty and student work and research for accreditation purposes, academic awards, publication information and curriculum impact.

For more information on ePortfolios, go to: http://at.tccs.tufts.edu/