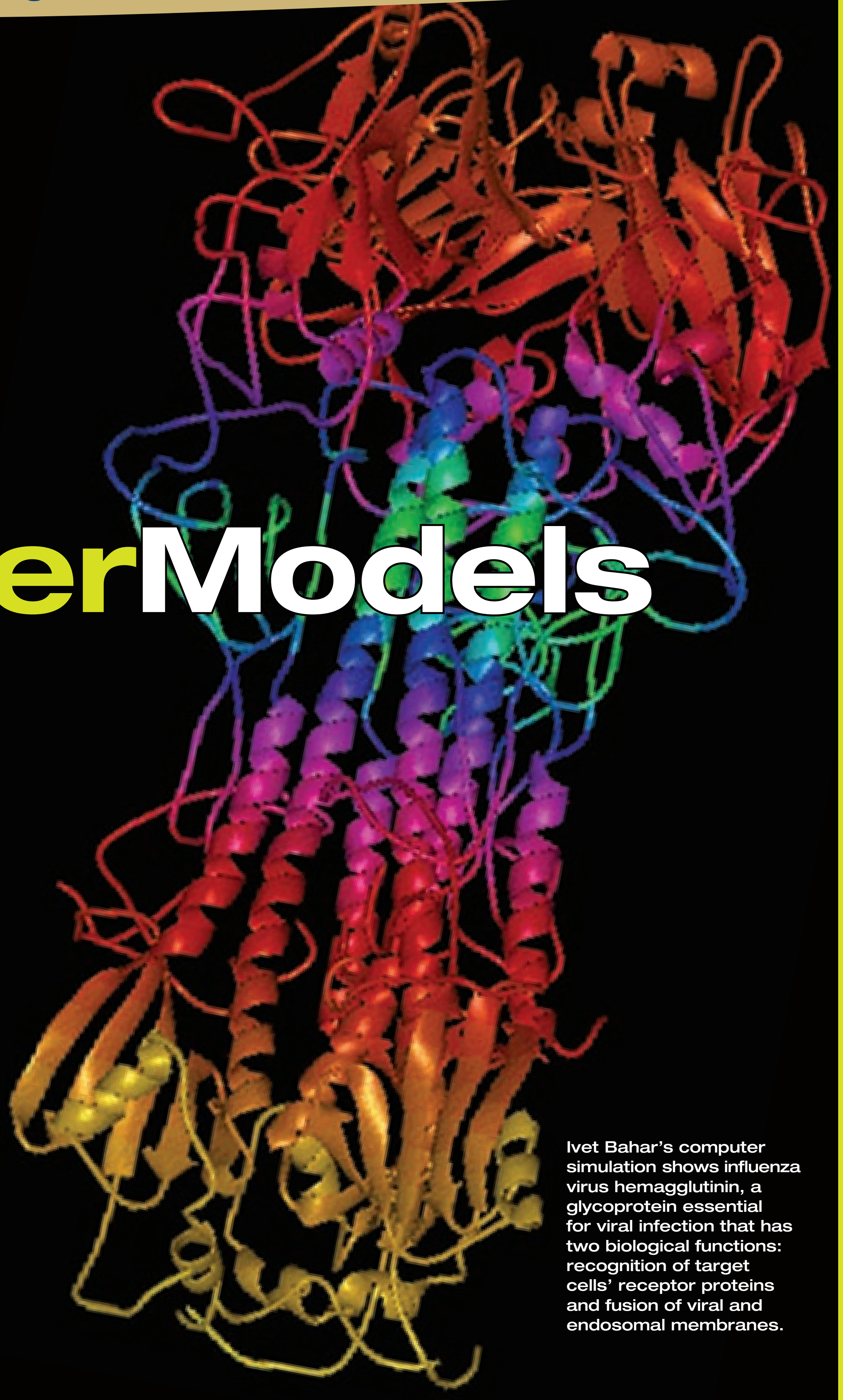
SuperModels



Ivet Bahar's computer simulation shows influenza virus hemagglutinin, a glycoprotein essential for viral infection that has two biological functions: recognition of target cells' receptor proteins and fusion of viral and endosomal membranes.

University of Pittsburgh researchers are crafting complex computer models—ranging from water molecules and the economy to folding proteins, nanotubes, and the cosmos—that are taking on a whole universe of questions, literally, and positioning Pitt as a major center of computational simulation.

You think you know how water works: put it in the freezer, it turns to ice; put it on the stove, you get steam. Kenneth Jordan knows, however, that water is anything *but* simple. University of Pittsburgh professor and chair of the Department of Chemistry and director of Pitt's Center for Molecular and Materials Simulations, Jordan and his research group are studying the behavior of very small amounts of water—containing six to 30 molecules of H_2O.

The wet sextets can form any of several hundred different shapes, from hexagons to prisms to open-book-like forms. Changes in energy can drive the little groups into different arrangements, and some shapes perform better than others in capturing stray electrons. Understanding water, Jordan contends, is a key to fathoming how life itself works—and how it sometimes falters. Water's interaction with electrons, for instance, could be a factor in the process by which radiation damages DNA.

But in a lab, you can only do so much with six water molecules.

Bring on a computer system that can simulate complex behaviors, however, and the research possibilities expand exponentially. Indeed, welcome to Jordan's world of discovery, where traditional, hands-on experimentation has converged with the virtuality of sophisticated computer modeling and simulation. With his trusty rack of 32 central processing units (CPUs) connected by an Ethernet switch, Jordan can simulate the complex behavior of water molecules under billions of conditions—all without getting his shirt wet. He then can compare notes with collaborators at Yale University, who are performing experiments with water.

Says Jordan of this combination of simulation and experiment: "It gives a much clearer picture than either the experiment or the [computer] model can provide."

Hard to imagine six water molecules? Let's think bigger. Proteins are molecules with thousands of atoms. The U.S. economy has all the diversity and volatility of an ocean. And the universe—well, things don't get any bigger than that. All of the above are systems, though, and they're all the

subjects of ongoing computer modeling and simulation efforts by Pitt researchers. By creating virtual worlds, researchers are advancing our knowledge of the real world in ways that may eventually make today’s understanding of chemistry, biology, economics, and astronomy look like a water droplet in the bucket.

At the same time, these researchers are bringing international renown to Pitt for their efforts.

Faster, Smaller, Cheaper

Jordan says his journey from lab chemist to computer whiz began in 1974, after a Carnegie Mellon University chemist named John Pople released a program for understanding the locations of electrons in molecules. Jordan installed the program at Yale, where he was an instructor at the time. But he wanted to be more than a downloader of other people’s codes.

“There are those folks for whom the computer is simply a tool—the software gets dumped on it, and you use it. And there are others who have been seduced by it,” says Jordan.

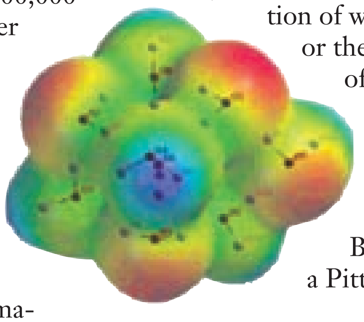
Count Jordan among the seduced.

The computational power of the cluster of CPUs currently just a few steps from Jordan’s office was unimaginable in 1970. Even 20 years ago, it took a mainframe computer costing $300,000 to do one million calculations per second, he says. His cluster cost $30,000 and can perform 50 billion calculations per second. And the top computer at the Pittsburgh Supercomputing Center (PSC), just up the Parkway in Monroeville, can do 6 trillion math operations per second—though that machine, known as Lemieux, can’t be called cheap. It cost $36 million.

The increased speed, reduced price, and improved storage capacity of computers have spurred a revolution in computational modeling and simulation at Pitt, making it a leader in the field. Of course, modeling takes many forms, and a model for simulating the behavior of six water molecules obviously is different from one that mimics galaxy formation. But the principles are roughly the same.

Typically, researchers observe the behavior of a system and reduce everything that’s known about its tendencies to repeating equations called algorithms. They input values for certain variables—say, the temperature of the collection of water molecules or the expansion speed of the universe—and tell the computer to play out different scenarios.

Bard Ermentrout, a Pitt professor of computational biology and mathematics, has written programs that model systems as diverse as ant colonies, the neurons involved in processing visual images, and the internal mechanics of cells. To Ermentrout, computer models serve three purposes. First, they allow him to prove or disprove theories about how things in the real world actually work. Second, they provide a way to test a system’s reactions to conditions that would be difficult or impossible to create in a lab.



This model is color-coded to show the positive and negative regions of nine water molecules. Blue represents a positive charge and red respresents a negative charge.

“Finally,” Ermentrout says, “they suggest new experiments and fresh ways to look at old data.”

The interplay between modeling and experimentation is critical, he contends. If the computer’s results don’t match what’s seen in reality, researchers tweak the algorithm until the two come closer. Once they match, researchers can postulate that the equations they’ve designed mathematically describe the system’s real-world behavior.

Carson Chow, an associate professor of mathematics and neurobiology at Pitt, for instance, is modeling the onset of sepsis—an often-deadly overreaction of the body’s immune system in which inflammation damages otherwise healthy tissue and begets more inflammation. Chow says that no effective treatment for sepsis currently exists.

“The [sepsis] process is extremely complicated and involves many types of immune cells and signaling hormones,” he says.

So Chow and collaborators are building a virtual human immune system at the same time that they test the response to sepsis in mice and rats.

A Notable Advantage

Researchers around the world are involved in computer modeling, but those at Pitt have a notable advantage. Pitt is a partner in PSC with Carnegie Mellon University and Westinghouse Electric Company. That partnership affords Pitt faculty relatively easy access to Lemieux—the world’s second fastest computer at its introduction—and several other very powerful computers. More than 30 Pitt researchers have made use of the supercomputing center’s machines, says Ralph Roskies, codirector of PSC and a physics professor at Pitt.

University researchers seeking to use PSC’s machines go through the same process as do scientists at any academic institution nationwide. They apply, and if their projects warrant use of the center’s computers, they get time on the machines, gratis. Pitt’s position as a partner in PSC confers other benefits.

First, PSC often collaborates with Pitt scientists on projects, and its imprimatur adds significant heft to any grant application. Second, Pitt researchers often are present at seminars hosted by PSC, affording them opportunities to network and share their findings with researchers from around the world. Third, the proximity of PSC to Pitt and Carnegie Mellon allows personal contact with PSC consultants. Fourth, PSC staff can provide expertise for computing efforts at Pitt. And fifth, PSC is one of several Pitt-Carnegie Mellon projects that encourage collaboration between researchers at each institution.

Says Roskies: “We act, if you wish, as the glue that sometimes connects people from both universities.”

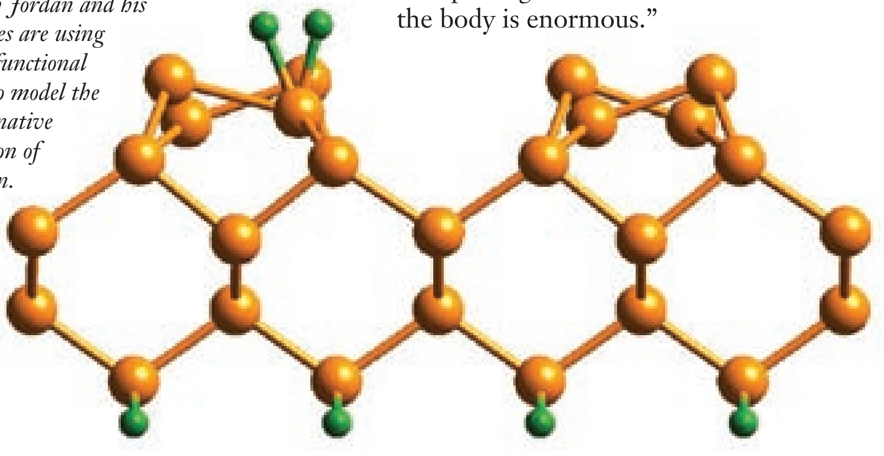
Jordan doesn’t need Lemieux to run his simulation of electron capture by a collection of six water molecules. His algorithm simulating the behavior of the six molecules occupies just eight of his interconnected CPUs—for about a week. But he has worked on other projects at PSC. Jordan and J. Karl Johnson, associate professor of chemical and petroleum engineering at Pitt, are working on joint projects with the computational group at the South Park–based National Energy Technology Lab. Pitt researchers also collaborate with scientists at the University of Pittsburgh Medical Center, Duquesne University, and such private companies as IBM, in addition to those at Carnegie Mellon. Together, these partnerships comprise one of the largest concentrations of computer modeling work in the world. And since many of the partner organizations are intimately involved in solving real-world problems, the potential results aren’t just academic.

For instance, if the thought of a billion calculations a second gives you a migraine, Rob Coalson may have the model for you. Coalson, a Pitt professor of chemistry and physics, and his team write programs that mimic one of the body’s least understood, but most important, structures: ion channels. The microscopic channels carry such charged atoms as hydrogen, potassium, and chlorine through the body and are essential for everything from the creation and propagation of nerve signals and muscle contraction to the production of such chemicals as insulin.

“Malfunction of ion channels is implicated as the root cause of many serious diseases and neuropsychiatric disorders, including cystic fibrosis, certain types of epilepsy, and migraine headaches,” Coalson says.

According to Coalson, it’s difficult to test ion channels in their natural environment—within a living organism—but it’s important that they be better understood, since about half of the drugs on the market today target ion channels. Coalson’s group has developed simulations that calculate the rate of flow of ions through channels in cell membranes, under various voltages and electrolyte concentrations, in an effort to gain insight into the relationship between channel structure and function.

“Clearly,” Coalson says, “the potential impact of understanding how ion channels work and of correcting or replacing defective ion channels in the body is enormous.”



Kenneth Jordan and his colleagues are using density functional theory to model the recombinative desorption of hydrogen.

Biology: Unfolding Proteins

Ivet Bahar, a professor of molecular genetics and biochemistry in Pitt's School of Medicine, takes off her glasses and moves the hinged arms in and out. She demonstrates that her frames are supposed to bend in only two places, but that a pair of specs offers infinite variations of position between a fully opened and fully closed pair.

Proteins are similar, says Bahar, who also is director of the University's Center for Computational Biology and Bioinformatics. The complex molecules have one or more "hinge sites," which can flex to different angles and lengths, allowing the "nodes" of atoms to which the hinges connect to bond to other molecules and perform their functions. The hinges are points of vulnerability *and* opportunity.

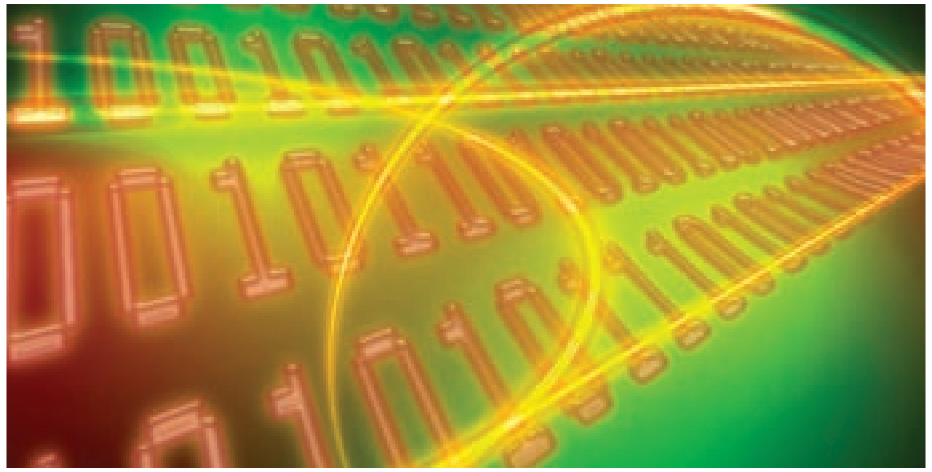
"We can say that, in this molecule, this [hinge] site is critical, and a mutation here may be disruptive to function," says Bahar. On the other hand, a particular hinge may be the ideal bonding spot for a therapeutic drug.

Since 1995, Bahar's team has been perfecting its elastic network model of proteins. Since proteins can contain thousands of atoms that can form billions of different configurations, the task of tracking the movements of every atom would challenge even Lemieux's considerable computational power. But by concentrating on the nodes and springs, Bahar's computer model can analyze the likely behavior of a medium-size protein in about five minutes or a really big protein complex in about two days.

Her team has run nearly all of the 23,000 proteins with known molecular structures through the elastic network model. They also have compared the model's predictions to experimental data on the behaviors of an estimated 100 proteins, and so far the virtual and real-world results match up nicely.

"Now we know what these structures are likely to do," Bahar says. "We think this is a very useful tool for understanding how proteins work."

The study of how proteins work is often called functional genomics, and it's a fledgling science that likely will depend heavily on computer modeling. The cracking of DNA's code, says Bahar, was an important step, but it also revealed that much of life's functioning was governed by the actions of tens of thousands of proteins. So genomics spawned structural genomics, and, more recently, proteomics, which seeks to map the structures and interactions of complete sets of protein molecules.



"Knowing structure is not enough," says Bahar. "We want to understand what it does." So structural genomics then led to functional genomics. But with so many proteins, Bahar says it may be impractical to conduct expensive laboratory studies on every one. However, with the elastic network model, researchers can gain a high level of understanding of a protein's function just by inputting the structure and letting the computer work its seeming magic.

The elastic network model currently is available online for researchers anywhere to use. Bahar says she thinks pharmaceutical companies soon will begin using the model, too. Eventually, her work and that of other functional geneticists could lead to the development of the ultimate medicines.

> "There are those folks for whom the computer is simply a tool—the software gets dumped on it, and you use it. And there are others who have been seduced by it."
>
> — *Kenneth Jordan*

"We want to be able to design proteins that have specific functions," she says.

John Rosenberg, a professor of biological sciences at Pitt, agrees. "Someday we hope to be able to really manipulate proteins in a directed way for practical purposes," he says. "That's going to be very hard if we don't know how they work."

Whereas Bahar's models analyze many proteins, Rosenberg has concentrated his efforts on just one: a bacterial enzyme called Eco R1 endonuclease, which consists of some 5,000 atoms. In his simulations, various stimuli "perturb" the Eco R1, and the computer predicts what the enzyme's 5,000 atoms, plus 55,000 other atoms in the surrounding water molecules, will do.

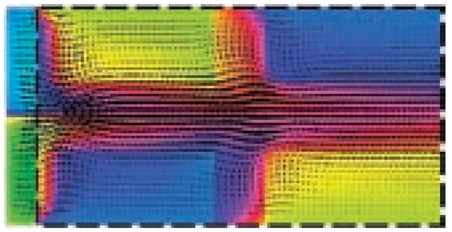
Eco R1 has the ability to slice DNA with great precision, and Rosenberg says its function in nature appears to be to attack invading DNA, like that of invading viruses. In modern biology, though, the enzyme has another role: cutting DNA that can then be spliced back together into a more useful, or healthful, configuration.

"Cloning," he notes, "would not be possible without precise tools like this."

Chemistry: Little Reactions

From the advent of the calculator to the rollout of the Palm Pilot, mathematicians and computer scientists have had a monopoly on handheld devices. Chemists, meanwhile, have had to do much of their work in room-sized labs. That's changing quickly at Pitt. The ongoing effort to build "microreactors," or "labs on a chip," is moving Pitt chemistry away from the black Formica table and into the palms of the researchers' hands. Miniaturization might not allow chemists to do experiments while waiting in the lunch line, but it could enable them to use much smaller samples of expensive materials—and that could be especially important in drug discovery.

One problem, though, has been how to mix fluids that don't want to combine, like oil and water, in the hairsbreadth tubes called microchannels that are the "circuitry" of a microreactor. Anna Balazs, Pitt's Robert Von der Luft Professor of Chemical Engineering, and Research Associate Olga Kuksenok, also in Pitt's Department of Chemical and Petroleum Engineering, decided to address that issue by taking their cues from nature. The exteriors of some organisms, from desert beetles to tropical plants, are really intricate mosaics of different surfaces that attract and repel water. Together, the hydrophilic and hydrophobic surfaces direct the water to the point most advantageous to the organism—say, the beetle's mouth or the plant's base. Build a similar mosaic of surfaces that attract different fluids inside a microchannel, arrange the surfaces so they pull two liquids back and forth, and you'd have the perfect little blender, the scientists contend.



Computer models developed by Anna Balazs' research team have shown that creating a checkerboard pattern of water-attracting and repelling surfaces inside hair-thin microchannels enables materials passing through the microchannels to mix.

In John Rosenberg's laboratory, the structure of Eco R1 endonuclease, a bacterial enzyme which can be used to precisely cut through DNA, is revealed through X-ray crystallography.

But what's the optimal arrangement of the materials on the inside surface of the microchannel? And precisely how do the fluids mix? Those are the kinds of questions a computer model can best answer. So Balazs, Kuksenok, Pitt Physics Professor David Jasnow, and collaborators at Oxford University in England crafted an algorithm describing what's known about the behaviors of immiscible fluids—those that don't want to mix together—and attractive surfaces. They found that a checkerboard pattern of surfaces inside the microchannel could optimally pull the fluids back and forth through each other, causing the most possible mixing. And their model allowed them to input the viscosities of any two liquids, the different levels of attraction between the liquids and the surface materials, and the flow rates of the liquids, and produce a color-coded video showing exactly how the liquids are behaving in the microchannel.

"What was unique about these studies was coupling these chemical patches and the imposed flow," says Balazs. "While it has beautiful physics involved, it also has practical applications, too."

For instance, if you want to precisely control the flow of an oil-soluble drug into a patient, you have to know exactly how that drug is mixing with the water base that carries it into the bloodstream. With their model, you can see exactly how the liquids are interacting, and where greater concentrations of one or the other exist. An Oxford researcher now is trying to reproduce the results in a lab.

Microchannels may be perfect for mixing fluids, but to purify some gases, you need something even smaller: a nanotube.

J. Karl Johnson, associate professor in Pitt's Department of Chemical

These functional forms of hemoglobin demonstrate the utility of computer simulations, here from Ivet Bahar's laboratory, in predicting the changes in protein structure associated with biological function.

and Petroleum Engineering and also a B. P. America Faculty Fellow, is using molecular modeling to help reduce the cost of purifying gases by pumping them through carbon nanotubes, which are like tiny straws one billionth of a meter in diameter. Because the tubes are incredibly smooth, the gas can move as much as 1,000 times faster than through conventional separation membranes, while the carbon purifies it. Perfecting the process in a computer model, he says, is a cost-effective prelude to lab work. Success could revolutionize the technology involved in gas masks, environmental filters, and even power plant emissions-control systems.

Balazs' group of researchers also has delved into the growing field of nanotechnology. They have created computer simulations of the interactions of polymers and tiny inorganic nanoparticles, in the hopes of finding a way to coax the molecules into assembling themselves into films that might be used in the future generations of batteries or photovoltaic devices. Those devices would have many uses; they might even power the aforementioned handheld chemistry labs, which industry analysts say could become a multibillion-dollar market.

Economics: Revealing Tension

Industry analysts make lots of predictions, and sometimes it seems they're wrong as often as they are right. That's because the economy is a gigantic equation with many variables, including the rate of technological advance, consumer confidence, the flow of information, and geopolitical stability. Despite the economy's complexity and susceptibility to seemingly random events, its overall performance *can* be modeled, says Pitt Department of Economics Chair Jean-François Richard.

Working with Roman Liesenfeld at Kiel University in Germany and Pitt colleague David DeJong, a professor of economics, Richard and his group have analyzed changes in U.S. Gross Domestic Product (GDP) since 1950. Richard and DeJong have found that the U.S. economy has a "sustainable" growth rate of 2.5 percent per year. But it never settles in at that rate. Instead, Richard claims, it bounces around like a ball attached by a rubber band to the wheel of a slowly climbing train in gusting winds. The tautness of the rubber band is the key to predicting what the ball will do next, he says.

"Deviations from the sustainable growth rate will create tension," says Richard, and eventually the invisible band pulls the economy back the other way.

Their study of past GDP changes led Richard and DeJong's team to develop a calculation called the "tension index," which is a function of the economy's distance from its sustainable growth rate.

"The tension index is a very good indication of when the economy will turn up or down," Richard says. "The higher the tension index, the more likely [the economy's direction] is to reverse."

The tension index is the centerpiece of Richard's new GDP Forecasting Model. The model can determine the likelihood that the economy will shift from contraction to expansion mode, or vice versa. With somewhat less precision, it also can predict the velocity of an economic rebound or downturn, once that change in direction has occurred.

"It has become a little more difficult than it was in the past," he notes. "There have been unexpected events like [the] September 11 [terrorist attacks] and the war [in Iraq]."

So Richard's team runs a number of different scenarios and produces a full disposition of possible GDP growth paths that extends two years into the future.

Despite the economy's complexity, its overall performance can be modeled, says Jean-François Richard.

"We do better than any existing model of GDP forecast," says Richard. "It has already drawn a lot of attention in the academic community, because predicting GDP is of great interest to macroeconomists."

Richard says he eventually plans to add into his model such economic factors as unemployment levels, stock market fluctuations, and fuel prices. Their inclusion could lead to even more precise predictions of growth. In the meantime, he's unveiling the GDP Forecast Model to the press and the wider world.

Will the model's results hold up under scrutiny? Richard says he thinks so, in spite of the uncertainty of the times.

Astrophysics: Unlocking the Cosmos

It will take a while to see for certain how well Ravi Sheth's models work. But within a few billion years, it probably will be obvious whether the universe is expanding at an ever-accelerating rate or at a decelerating rate or whether it is beginning to collapse.

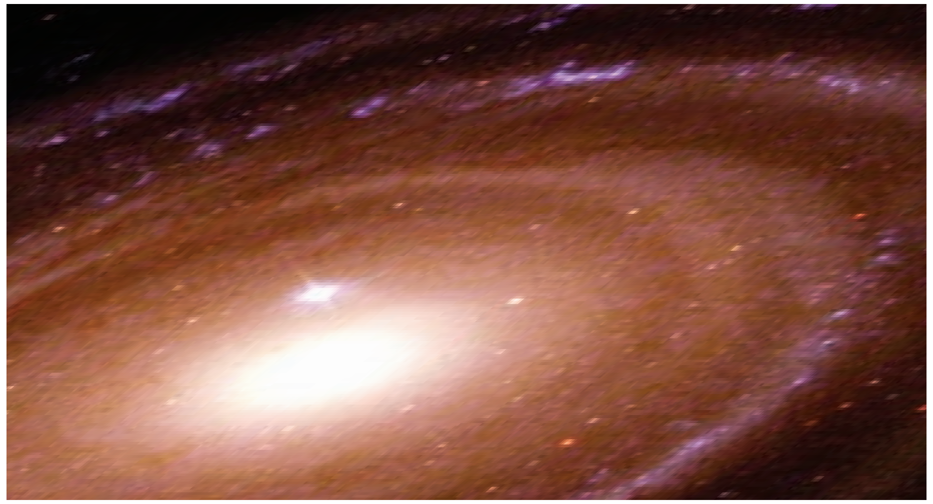
That's the central question in astrophysics these days, and researchers Sheth and Andrew Connolly, both assistant professors in Pitt's Department of Physics and Astronomy, are at the center of efforts to answer it. Pitt—with active participation from Sheth and Connolly—is one of 13 institutions of learning worldwide that are participating in the Sloan Digital Sky Survey, an effort to map one quarter of the observable universe, including some 100 million celestial objects.

The result of the Sloan survey, Sheth says, will be a database with nearly the same amount of information in it as the entire Library of Congress. Connolly is focusing on mining that data so scientists can draw useful conclusions about the 99 percent of the objects that are so dim or so distant that their full light spectra haven't yet been recorded. Sheth says he is looking at the incredible variety of viewable objects and comparing them to the results of computer simulations.

"There are so many objects [in the Sloan database] that you can now start asking questions about which ones are weird—the kinds of things that might be one-in-a-million but might disprove existing theories," Sheth says.

Those existing theories hold that the universe was an almost perfectly uniform, incredibly dense, and rapidly expanding soup of electrons, protons, and light roughly 300,000 years after the so-called "Big Bang." That's when things cooled enough for hydrogen to form. Meanwhile, as the theories go, tiny fluctuations in density caused some particles of newborn matter to pull together. In computer simulations, Sheth and other astrophysicists start at that point in time and create a virtual universe of 16 million or 100 million particles. (The biggest simulation to date involved 1 billion virtual particles, each representing approximately the mass of five galaxies.) They plug in a rate of expansion, factor in the gravitational force of each particle on every other particle, and see what happens.

What happens, Sheth says, is that, even as the whole universe expands, the particles pull together into web-like filaments and then flow along those filaments to intersections, where they eventually form clusters of galaxies.



Determining whether the universe will collapse into unimaginable density or expand into cold blackness may prove the ultimate test for comupter modeling. Work by Pitt researchers Ravi Sheth, Andrew Connolly, and their colleagues will likely advance the science of computer modeling as much as it advances astrophysics.

"You start with something smooth, and it becomes lumpy," he explains.

The exact patterns that result depend on the rate of expansion plugged into the model, whether that rate accelerates or decelerates, and the amount of matter postulated. Currently, Sheth and colleagues Joerg Colberg, a research associate in Pitt's Department of Physics and Astronomy, and Jeff Gardner, an adjunct professor in the department, are using the supercomputing center to run a 100 million-particle simulation of the universe, in which each particle represents the mass of about 50 galaxies. They're also analyzing other peoples' models and comparing the predictions to the images provided by the Sloan survey. The work would have been cumbersome, if not impossible, just a few years ago, when it took supercomputers three weeks to run a 16 million-particle simulation. Now that same simulation takes only a few days on a decent desktop computer.

Understanding the expansion rate of the universe won't stop DNA damage, create new drugs, help enable handheld chemistry labs, or unlock the mysteries of the business cycle. But determining whether the universe will collapse into unimaginable density or expand into cold blackness may prove the ultimate test for computer modeling. As such, the effort at Pitt to unlock the riddles of the cosmos may advance the science of computer modeling as much as it advances astrophysics. Sheth says he studies the universe in part to prove "there can be models that work."

Adds Sheth: "It's more than just 'because it's there.' Einstein said that the most incomprehensible thing about the universe is that it is comprehensible. The pleasure is in finding models which bear this out." ■