Columbia Engineering can claim that the information era began when Herman Hollerith, an 1879 graduate of the Columbia School of Mines, founded the company that was to become IBM. Today, computers, microcomputers, computerized machinery, robots, fiber optics, and all manner of digital technologies provide a research area in which many of our faculty are engaged, by advancing digital frontiers and cybersecurity to keep our information safe.
For Professor Alfred Aho, the question is simple: “How can we get reliable software from unreliable programmers?” The issue is more than academic. Aho can point to such high-profile fiascos as a $1 billion write-off for failed flight control software and hundreds of millions of dollars spent fixing an airport’s automated baggage handling system.

In fact, a 2002 National Institute of Standards and Technology (NIST) study found that software defects cost the economy $60 billion annually and account for 80 percent of software development costs. Even then, Aho estimates that most commercial software has 1,000 to 10,000 defects per million lines of code.

“If you’re developing a computer game, that doesn’t matter much. But if you’re programming a pacemaker, it’s a matter of life and death,” he said.

Aho’s goal is to create a system that automatically tags potential problems. He hopes to do this by using the technology behind compilers, programs that translate easy-to-use programming languages like C into instructions a computer processor can understand.

When a compiler translates a program, it captures details about how it was built. Aho wants to compare this actual implementation with the program’s technical specifications, which define such things as naming conventions, allowable operations, associations, data sets, and order of functions. This is similar to inspecting a building’s structure, wiring, and plumbing against schematics and code.

Software, however, is more complex. “Let’s say the source program has 1 million lines of code, and you want to look for all examples of addition,” Aho explained. “It’s written by several different people. Some may not have consulted the specification. They might use their own names for variables. Instead of writing ‘add,’ they might write it as ‘plus.’”

Those subtle changes make it incredibly difficult to track errors. A plus function might use different data types than an add function, and produce unequal results. Or a programmer may discover a problem involving add functions, but fail to look for plus functions to see if the same problem exists.

“All large programs have a specification document that itemizes how the program should be written. I would like to specify a property from this document and test for its properties in the software. We already know how to create tools that do some of this in compilers. Now we want to extend these tools to software development,” Aho said. “This is a long-term project, but if we can make a small dent in software development and maintenance costs, we can save billions of dollars.”
Could centuries-old techniques used to classify species hold the key to computerized face recognition? Professor Peter Belhumeur certainly thinks so. Face recognition has many potential uses, from verifying financial transactions to recognizing terrorists and criminals. Today’s systems work by superimposing a subject’s face over images in a database. If they align, the computer samples pixels from each image to see if they match.

The process is not very reliable. “Recognition algorithms make mistakes that they should never make, like confusing men with women, or one ethnicity with another,” Belhumeur said. Belhumeur was working on improving those algorithms when Smithsonian Institute taxonomists asked for help developing software to classify plant species from photos of their leaves.

Instead of superimposing images or matching pixels, Belhumeur drew on the wisdom of taxonomists dating back centuries. They classified plants by asking a series of questions whose yes-or-no answers narrowed the choices until they came to the right plant.

Belhumeur’s software hopes to do this automatically. “We want to develop software that can determine whether a leaf is simple or compound, lobed or unlobed, and smooth or serrated along its edge, and then use these determinations to identify the species,” he said. “This is exactly the opposite of how computerized object recognition is done. Instead of pixels, we want to compare visual attributes,” he explained.

Belhumeur wondered if he could use a similar strategy to recognize faces. “Could we develop software that made qualitative decisions about each image? Is it a male or female? Young or old? Broad or pointy nose? Facial hair or not? If we could build reliable classifiers to answer these questions, we could search for pictures based on their attributes,” he said.

Belhumeur’s system uses roughly 100 labels, ranging from eye and nose shape to hair color and gender. In tests that compare a photo to a known image, like an identity card, it outperforms pixel-based technologies, he said.

It also makes it possible to search for pictures with words that describe visual attributes. “We could search through a database based on a victim’s description of an assailant, or use it to search one’s seemingly endless collection of digital photos,” he concluded.

“Borrowing a concept from taxonomy, we could simplify object recognition by labeling objects in images with a list of highly descriptive visual attributions. A computer could then classify an image by recognizing these objects. We might also use text labels to browse or search for pictures.”

Accurately searching through the glut of visual data available today—digital images produced daily in the thousands or millions—depends upon how closely your verbal description matches the words used to classify the image. It is a frustrating, time-consuming exercise that affects the general user as well as news, media, government, and biomedicine specialists who crave a richer search and browsing experience. An automated visual matching and search technology would not only enhance classification and searching activities, but could also facilitate media forensics, helping to explain if an image has been manipulated, or if it is a natural photograph or computer graphic.

In order for a search engine to visually classify and find images, or determine if images have been tampered with, computers would need to perceive the abundant visual information provided by each individual image. All that data is much like DNA: thousands of genetic concepts of objects, people, scenes, events, and domain-related syntax that make up the individual image.

Shih-Fu Chang, professor of electrical engineering and director of the Digital Video and Multimedia (DVMM) lab at Columbia Engineering, targets his research on next generation search engines for digital images and videos, and has been influential in shaping the vibrant field of content-based multimedia retrieval. His group leads the University’s ADVENT University-Industry Research consortium, promoting industrial collaborations with Columbia’s research teams in the media technology area. In addition, his group has actively participated in the development of MPEG-7 and MPEG-21 international standards.

Ranked by Microsoft Academic Search as the most influential researcher in the field of multimedia, his research includes multimedia search, pattern recognition, media analytics, video adaptation, and media forensics. Results include a groundbreaking search paradigm and prototype tools that allow users to find content of similar visual attributes, search videos by a very large pool of visual concept classifiers, and summarize the event patterns and anomalies found in a large collection of video content.

In 1998, he developed one of the first video object search systems, VideoQ, which supported automated spatio-temporal indexing at the object region level. His work has been broadly funded by government and industry and many video indexing technologies developed by his group have been licensed to companies.

Chang is a Fellow of the Institute of Electrical and Electronic Engineering (IEEE) and received the IEEE Kiyo Tomiyasu technical field award in 2009.

B.S., National Taiwan University, 1985; M.S., California (Berkeley), 1991; Ph.D., 1993
In many ways, a good theory behaves like a rock thrown in a pond: It makes a splash and then its ripples spread. Associate Professor Maria Chudnovsky's work in graph theory is like that. "A graph is a good model for many practical problems, a good way you can think about them. You can think of the Internet as a graph and the computers on it as vertices; some are connected and some are not. Graph theory can tell us about its structure."

Graph theory does not involve what we normally think of as graphs. Instead, it involves groups of points, or vertices. Sometimes they form geometric objects like squares and pentagons. Other times, they are distributed as randomly as cities or cell phone towers on a map.

Graphs are characterized by the properties of their vertices and the lines, or edges, between them. They can be used to answer problems, from finding the best route for a delivery truck to routing Internet traffic to calculating the shortest itinerary on a GPS.

Chudnovsky works at understanding these attributes. In 2002, her team proved a conjecture about perfect graphs, which are graphs roughly defined as being easy to color. They showed that only two types of defects keep a graph from being perfect, and that all perfect graphs fall into a handful of different categories.

Chudnovsky's proof makes it possible to determine if a graph is perfect without coloring all its vertices. While this may sound like a strictly cerebral exercise, perfect graphs were originally conceived in order to solve a problem in communications theory.

Her work is relevant in other fields as well. Engineers could use her proof to locate wireless towers so their frequencies do not interfere with one another. Knowing whether a graph is perfect or not also helps computer scientists choose efficient algorithms to solve certain problems.

Chudnovsky continues to explore the structure of graphs. Her recent work looked at graphs that did not contain a claw. This structure occurs where three lines, or edges, emanate from a common vertex to form a three-fingered claw.

"We've explicitly described all graphs that do not contain a claw. Now that our characterization is in place, many problems that seemed to be out of reach can be solved relatively easily," she said.

While her work is highly abstract, her results promise to solve some of the most practical of problems.

Emanuel Derman knows something about models. He practiced physics after receiving his Ph.D., but moved to Wall Street in 1985. At Goldman Sachs, he co-developed one of the earliest interest rate models, and later headed their quantitative strategies group. Business Week chose his memoir, "My Life as a Quant: Reflections on Physics and Finance," as one of the top 10 books of 2004.

"Models Behaving Badly" is Derman's tentative title for his next book. "It's about the different approaches people use to understand the behavior of the world," he said. In it, he distinguishes how theories differ from models, and explains how the unwarranted assumptions of models can lead to incorrect conclusions.

"Theories," Derman explained, "are attempts to grasp the way the world actually is, even if we don't know why. Take Newton's laws. You can't ask why they are correct. That's the way the world is. These are regularities that are always true."

Models are different. "In my view, they are metaphors or analogies," Derman continued. "We say, 'The brain is like a computer,' or 'Stock prices change the way smoke diffuses through a room.' Models are attempts to describe something by using theories that already work in a different field."

"When I first came to finance, I used the principles of physics to try to build something just as truthful. I discovered that although the techniques appear similar, the resemblance is deceptive. When we make analogies, we simplify things," he said.

Many on Wall Street believed their models represented reality. They were disabused of that notion in 2008.

"In physics there may one day be a theory of everything. In finance and the social sciences, you’re lucky if there is a usable theory of anything."

Yet models still have a role to play. "I'm a bit of a Platonist," Derman said. "I think there is some truth out there. I'm trying to distinguish between finding the truth, which is rare, and building models while understanding their inherent limitations."

"You know, Maxwell once remarked that Ampere's experiments could not have led to his results. His experiments seemed to confirm his intuition rather than point to it."

"I believe in intuitive knowledge, but you don't just wake up with it. It comes after a lot of hard work. Models are a step on that road," he concluded.

B.Sc., University of Cape Town, 1965; M.A., Columbia, 1968; Ph.D., 1973
There is a big difference between hearing and listening. Listening requires complex auditory processing, which facilitates learning. It’s a skill humans use automatically in order to filter out background noise to understand someone’s speech; remember a previously-heard tune and hum along; or recognize the difference between a ringing phone and ringing alarm and understand what an appropriate response to those sounds would be.

Human listeners are able to handle such mixed signals, but machines—such as automatic speech recognizers—are vulnerable to added interference, even at levels that listeners barely notice. Consider the implications of machines that could respond when called, technology that could classify and retrieve videos by their sound tracks, or applications that could automatically search for audio data the same way we do now for text data.

To make these advances possible, it is important to understand how perceptual systems manage to make precise judgments in noisy and uncertain circumstances. This understanding can then be applied to extracting information from sound commonly encountered in daily life, identifying characteristics of the sounds, classifying them, and matching the sounds to appropriate responses.

Daniel P. Ellis, associate professor of electrical engineering, is working on such advances. He is the founder and principal investigator at the Laboratory for Recognition and Organization of Speech and Audio (LabROSA) at Columbia Engineering. This lab is the only one in the nation to combine research in speech recognition, music processing, sound separation, and content-based retrieval in order to implement sound processing in machines.

His chief focus is to develop and apply signal-processing and machine-learning techniques to extract high-level, perceptually relevant information from sound. His intention is to test theories about how human auditory perception works and enable the creation of machines that can make use of sound information in the same way humans do.

Ellis’ work in soundtrack classification pioneered the idea of using statistical classification of audio data for general classification of videos by their soundtracks. Current projects in the research group include speech processing and recognition; source separation and organization; music audio information extraction; personal audio organization; and marine mammal sound recognition.

He is a member of the Audio Engineering Society, International Speech Communications Association, Institute of Electrical and Electronic Engineering (IEEE) and the Acoustical Society of America.

B.A., Cambridge University, 1987; M.S., MIT, 1992; Ph.D., 1996

Delving into the Science of Listening

DANIEL P. ELLIS
Associate Professor of Electrical Engineering
The assurance of confidentiality is required in all aspects of transmitting information, from the exchange of banking information and health records to military tactics and trade secrets. The problem is there is no foolproof method to ensure that confidentiality. No matter how encrypted the information is that is transmitted, as long as there is a key to decrypt it, there is a weak security link in the chain of communication.

The solution may lie in quantum photonics, the sending and receiving of data in the form of photons – the tiniest particles that make up light. By sending data encoded in photons, the data stream becomes a single-use, self-destructing key. If the message is intercepted, the stream would change, immediately alerting the receivers to the breach. In addition, by intercepting the stream, the disturbance would automatically scramble the message, making it indecipherable.

Quantum photonic networks could decrypt classically-encoded messages in a matter of minutes – rather than months or years as per today’s networks. And such networks would allow for absolute security; even another quantum computer would not be able to secretly crack a coded message sent via a quantum network.

Working in the quantum world, addressing present day problems, requires an in-bred curiosity about the nature and behavior of matter and energy on the atomic and subatomic level, and a desire to develop revolutionary applications. Those are the talents of Dirk Englund, assistant professor of electrical engineering and of applied physics, who leads the Quantum Photonics Group at Columbia Engineering. He concentrates on quantum optics in photonic nanostructures, with primary applications in communications, computation, sensing, and energy. His research focuses on implementations consisting of quantum bits (qubits) that are encoded in photons and in spins of electrons and nuclei in semiconductors.

Englund’s work includes chip-based quantum networks that promise exponential speedups in computational algorithms and unconditionally secure cryptography as well as highly sensitive quantum-limited sensors. Recent works include time-resolved lasing action from single and coupled photonic crystal nanocavity array lasers, and optical modulation based on a single strongly-coupled quantum dot.

His group is also developing spin-off applications that rely on phenomena from cavity quantum electrodynamics (QED) to substantially lower the power consumption of optoelectronic systems for high-speed, low-power devices. These applications have potential for adaptation in high-performance computing. Related projects include radiation-hard electronics and radiation detectors and thin-film solar cells.

B.S., California Institute of Technology, 2002; M.S., Ph.D., Stanford, 2008
What happens when a mechanic must work on an unfamiliar piece of equipment? He or she will pull out a manual and keep referring to it while making repairs. Professor Steven Feiner has a better alternative, one that changes how we see the world around us.

His approach to this problem involves augmented reality (AR). Unlike virtual reality, which creates an artificial world, AR adds virtual information to the real world. AR can guide people through complex tasks. “Instead of looking at a separate manual while disassembling a PC, imagine putting on lightweight eyewear containing a see-through display that graphically highlights the screws in the order you need to remove them,” Feiner said.

Feiner has been developing experimental AR maintenance applications for 20 years. This involves delivering information about a system, quickly and naturally, as workers move around a workpiece. He does this by tracking the position and orientation of their eyewear, then aligning information with their perspective.

In recent studies with U.S. Marines at Aberdeen Proving Ground, Feiner’s lab found that AR helped professional mechanics find the location of parts they needed to repair faster than using manuals. “With manuals, the documentation is separate from the task. Workers are always going back and forth. AR keeps them focused on the work by integrating the documentation with the task,” Feiner explained.

Feiner is also working on better ways to display AR information for people interacting with their surroundings. He has come a long way since 1996, when his lab created the world’s first outdoor mobile AR system. Before the era of smart phones, ubiquitous GPS, and Wi-Fi, it consisted of head-worn and hand-held displays – plus a 45-pound backpack stuffed with electronics.

That system let users tour Columbia’s campus, overlaying the names and websites of academic departments on their buildings. Within a few years, Feiner’s lab had added multimedia news stories and created AR restaurant guides.

Today’s smart phones are far smaller and more powerful than those early AR systems. Feiner and his students are harnessing their power, both alone and with other computers and displays ranging from wearable to wall-sized.

AR displays can create a compelling experience. That is why Feiner wants to ensure that every AR system respects the physical environment and the user’s relationship to it. “We don’t want users losing awareness of the world around them while trying to cross a busy street,” he said.

B.A., Brown, 1973; Ph.D., 1987

Augmenting Reality

STEVEN K. FEINER

Professor of Computer Science

EXCELLENTIA COLUMBIA ENGINEERING
Computing power has grown rapidly, but not as fast as the problems researchers aspire to solve. “We’re dealing with enormous problems, problems so large we can’t even store all the numbers in computer memory at the same time. We cannot rely on the same methods we used for smaller problems and expect to solve them,” said Professor Donald Goldfarb.

His work on extracting movement from surveillance videos provides an example. On surveillance videos, the background never changes. One frame looks very much like the next, except for the people moving through the space. Each frame has roughly 20,000 pixels.

“To extract moving images from a couple of minutes of video, we need to process 50 million variables and 25 million linear equations,” Goldfarb said. Doing it by brute force—one computation after the other—would take days on powerful computers. Instead, he developed a systematic optimization procedure, or algorithm, that lets a simple workstation remove the background in under an hour.

Goldfarb has a long history of developing powerful optimization algorithms. Some of his early algorithms are used in commercial software to optimize complex systems. They make it possible, for example, to adjust refinery operations on the fly instead of spending weeks plotting a production schedule.

Goldfarb’s work goes beyond just finding fast ways to solve difficult problems. “I try to prove that the algorithms I develop are not just fast for a specific problem, but will work well for any similar problem. It’s like providing a certificate guaranteeing the algorithm’s performance,” he said.

He also tries to discover properties about different classes of algorithms. Recently, he has focused on convex functions. Like many algorithms, they recast algebraic problems in geometric terms in order to estimate answers more rapidly. Goldfarb likens a convex function to a bowl with the minimum, or optimal, value at the bottom. Constraints usually push the answers to any given problem somewhere along the sides of the bowls.

“If you’re sitting on the side, you can see every other point inside the bowl. If you look around and every other point is higher, then you are at the optimal point,” Goldfarb explained.

Recently, Goldfarb used convex functions to optimize a method to produce MRI and CT scan images using only one-fifth the radiation. “The algorithm enables us to get an appropriate image with fewer measurements, so patients only have to spend one-fifth as much time in these machines,” Goldfarb said.

B.Ch.E., Cornell, 1963; M.A., Princeton, 1965; Ph.D., 1966
Imagine searching for a concert and pulling up the usual web pages – plus untagged Flickr photographs, Twitter remarks, YouTube videos, and Facebook comments. Or asking when the band will perform again and getting back a table of dates and locations.

Associate Professor Luis Gravano is supercharging search engines to conduct exactly those types of searches. Often, that means tapping the chaos of social media. “It’s not so much about just returning a list of individual Web pages as it is about combining and making sense of all information on the Web to increase the effectiveness of a search,” he said.

For example, many online photos are tagged to refer to specific events. Others have time and GPS data that coincide with the time and location of an event. Sometimes photos are forwarded or linked to other people who have commented about an event.

“We analyze these tags, comments, and links, and automatically cluster them to correspond to real-world events,” Gravano said. His team has already shown that it can aggregate such information. It is now probing how to fit the data together to develop more powerful searches.

“If there is a concert or political demonstration, people take pictures, tweet, and form groups around these activities. We want to capture and associate this content with real-world events automatically. We’ll return results that correspond to a specific event at a certain time on a particular street in New York City.”

Gravano also wants to improve our ability to extract structured information, such as tables, from the Internet. Today, he explained, anyone who wants to analyze the characteristics of past infectious disease outbreaks would have to sift through hundreds or thousands of search engine results.

Gravano’s extraction technology searches for pages that are likely to contain the desired structured information, which is often embedded in natural language text. It then extracts, analyzes, and puts the information into a table automatically. Unfortunately, the process is prone to errors. Information is sometimes out of date or wrong. Writing is often ambiguous.

Gravano hopes to reduce errors by using such trusted sources as government documents, university archives, newspapers, and specialized websites, as well as by analyzing the frequency and context of the extracted information.

He also taps crowd wisdom to assess the reliability of popular sources. “Popularity is a step in the right direction – if you trust people to go to trustworthy sources,” Gravano said.

B.S., ESLAI (Argentina), 1991; M.S., Stanford, 1994; Ph.D., 1997
What do dresses, medical instruments, and the bristles on a paintbrush have in common? Their motion can all be predicted with unparalleled accuracy by techniques developed by Associate Professor Eitan Grinspun.

Grinspun’s techniques have broad application. In the movies, they produce stunningly realistic animations of gowns swirling on dancers and animal manes billowing in the wind. “If you can compute motions that obey physical laws, you can make artistic choices about what laws you want to disobey and produce things you would never see in real life,” he said.

His work is equally applicable to physics. “Think about how honey behaves when you pour it on a scone,” he said. “It is a liquid, but it loops around like a rope. If we can understand how honey moves, we can understand how lava flows or the best way to bottle shampoo.”

Bottling shampoo is not a trivial problem. Shampoo entrains air, which reduces its density and increases its volume. “If you can understand how shampoos move, you can reduce entrainment and pack them in smaller containers to reduce costs,” Grinspun said.

Physicians have used Grinspun’s techniques to test how to steer surgical needles through human tissue. Adobe has leveraged them to simulate each individual paintbrush bristle in its popular Photoshop and Illustrator programs. “Those bristles are really bending, and you get all the effects you would get with a real paintbrush,” Grinspun said.

What makes Grinspun’s work unique is his deep understanding of the geometry underlying physics. For example, when he looks at a long, thin surgical needle, he sees a flexible curve that bends and twists. “Computers, geometry, and physics are my ingredients. I mix them up in a bowl and what I get is a computer’s ability to predict the motion of materials.”

“We can visualize the problem by thinking of the boundary of North America on planet Earth,” explains Grinspun. “The energy stored in bending is like the continent’s perimeter, while the energy stored in twisting is its area. We have a competition between bending, which wants to keep length as short as possible, and twisting, which wants to deform the length to enclose more area.”

Understanding the geometry of those forces produces fast and accurate predictions of movement. The results are readily visible in movie special effects and in basic science as well.

B.A.Sc. in Engineering Science, Toronto (Ontario), 1997; M.S., California Institute of Technology, 2000; Ph.D., 2003
Professor Jonathan Gross knew little about Celtic knots before he started studying them. “I knew one when I saw it. They are characterized artistically by repetitive patterns and symmetries. Then, while browsing the Internet, I found a graphic artist’s description of them so precise, I could turn it into math.”

Gross uses computers to explore algebraic topology, the mathematics of translating geometric forms into algebraic expressions. “We calculate a polynomial from a picture of the knot. Once we represent the shapes with algebra, we can manipulate the math to learn fundamental truths about the shapes,” he explained.

“For example, Reidemeister proved that if you make new crossings in a knot without cutting the string, the resulting figure has the exact same polynomial as before. If you hand me a knot, I can either fumble for hours trying to untie it, or I can calculate a certain polynomial and quickly know that the string is really knotted,” Gross said.

Gross is quick to point out that his research is theoretical. Yet some of his insights have worked their way into practical applications. “Some of my work is related to practical technology, but what motivates me are mathematical problems that involve spatial visualization and deriving algebraic formulas to count mathematical objects far too numerous and/or too intricate to count by any elementary methods.”

Last year, for example, he collaborated with two colleagues in Texas to develop a computer graphics program to create designs in woven textiles. “We designed software whose mathematical models embody key principles of algebraic topology. A graphic artist doesn’t have to know any of this to use the software to create a complicated woven pattern very quickly,” Gross said.

Gross has also applied mathematical modeling to social anthropology. Anthropologists used to live with a people and describe what they saw. Their descriptions were typically highly subjective. Gross worked with a team that developed an objective way to measure and compare behavior.

They started with food systems. “There are differing levels of randomness in the way people eat. When some people eat scrambled eggs, you know for sure it’s breakfast. Not quite so for others. To differing extents, meal content reflects the time of day, time of year, and festivities. By measuring the information content in these patterns, we could make comparisons between different peoples,” Gross said.

To Gross, it was just another knot untangled by mathematics.

B.S., Massachusetts Institute of Technology, 1964; M.A., Dartmouth, 1966; Ph.D., 1968

Untying Knots with Mathematics
JONATHAN GROSS
Professor of Computer Science
Fill a glass half way and some people will call it half filled and others half empty. Either way, the amount of liquid in the glass is the same. Our frame of mind—optimistic or pessimistic—imposes meaning on what we see.

Investors in financial markets are not any different, said Assistant Professor Xue-dong He. They all view the same financial data, yet they draw different conclusions from what they see. Irrational biases often play a role.

“Classical economic theory assumes that investors evaluate information correctly and make decisions rationally. In reality, though, they have biases. They may overemphasize or overlook certain types of information, and this affects how they manage their portfolio,” He said.

For example, investors often miscalculate the odds of an event because they put too much weight on recent data. One common bias is to go with someone on a winning streak. “Gamblers who win two or three times in a row think they have a hot hand and are more likely to win the next time. The odds are still against them, but they over-emphasize their recent success,” He said.

Other investors may assess the odds correctly, but hidden biases guide their actions. “Look at people who buy lottery tickets and insurance,” he said. “A lottery ticket usually has an expected value lower than the selling price. Buyers know the probability of winning is very low, but take the risk for the reward. On the other hand, people know the probability of their house burning down is low, but they buy insurance because they are risk averse.”

He builds mathematical models that show how these twin engines—hope and fear—drive investment strategies. “Hope and fear coexist in investors’ minds. When stock prices surge, hope takes control, so investors are more likely to invest in stocks and gamble more. When the market turns down, fear dominates and investors quickly liquidate their portfolios,” he said.

“In financial engineering, not much work has been done on irrational biases. We have developed a concrete model of these irrationalities based on extensive research. We want to understand how these biases affect investor behavior and strategy,” he said.

He is one of the few financial engineers researching irrational motivations. By taking biases into account, he hopes to create models that better predict market behavior and perhaps even warn when investors are being carried away by irrational exuberance.

B.S., Peking University, 2005; D.Phil., Oxford, 2009

Modeling the Irrational

XUEDONG HE

Assistant Professor of Industrial Engineering and Operations Research
Anyone who has ever navigated an interactive voice recognition system to make a reservation or review a charge knows that anger and sarcasm change nothing. But one day they might, thanks to research by Professor Julia Hirschberg.

Hirschberg studies prosody, the intonation and melody of speech. Often, it conveys subtle differences in meaning. For example, “I like cats” may sound like a statement, but raising the pitch at the end turns it into a question.

“During deceptive speech, you experience emotions like fear if you think you’ll be detected or elation if you’re getting away with it. This shows up in the prosody of your speech. The best people at judging liars are criminals. Police were worse than average, and parole officers the worst of all, because they assume people are always lying,” she said.

Hirschberg’s goal is to teach computers to understand such subtle variations and reproduce them in natural sounding speech. This involves understanding how prosody changes under different circumstances.

“When I was at Bell Labs, we did lots of experiments that looked at people’s speech, and tried to predict what words he or she would emphasize,” Hirschberg related. “We looked at syntax, context, the part of speech being uttered – you use whatever information you have, and usually that’s not a whole lot.”

At Columbia, she has analyzed the prosody of charismatic and deceptive speech.

“Much of the perception of charisma is not about what people say, but how they say it,” she explained. “In English, charismatic speakers are very expressive, vary their pitch contour a lot, and speak more rapidly.”

She has also conducted extensive experiments in which people either lied or told the truth. In these experiments, the speakers told the truth about 61 percent of the time. Her automated computer system labeled identified truth tellers and liars about 70 percent of the time. Humans got it right about 58 percent of the time, worse than if they had just guessed “truth” every time, Hirschberg said.

She is also working on teaching interactive voice response systems a technique called entrainment. This occurs when one speaker mirrors back the same vocabulary, pitch and speed as another. “People like people who entrain to them more than those who do not,” Hirschberg said. “We want to teach computers to change their pitch, intensity, speaking rate and other factors to sound more like the user.”

If that doesn’t mollify the next generation of callers, at least the computer will recognize their anger when they express it.

B.A., Eckert College, 1968; Ph.D., Michigan, 1976; MSEE, Pennsylvania, 1982; Ph.D., 1985
Professor Garud Iyengar is helping to unlock the secrets of how colonies of bacteria work together, though he is not a biologist. "I’m a problem solver, rather than someone who focuses on one particular research area. My particular interest is in understanding how simple components can produce complex behavior when networked together," Iyengar said.

His varied background in mathematical modeling and optimization enables him to tease out insights that classically trained biologists might miss. "My particular strength is in building mathematical models to guide experimentation by blending tools, often from different disciplines, that together work better than any single tool used independently."

"Most scientists have a set of pet mathematical tools. Someone trained in statistics immediately thinks about regression to model experimental data. A computer scientist builds a combinatorial model. An electrical engineer wants to use information theory. I’ve been exposed to many of these disciplines, and so my bag of tricks is bigger."

Lately, Iyengar has been trying to discover how colonies of unicellular organisms communicate in order to exploit their environment. Density sensing in Pseudomonas aeruginosa, a bacteria that inhabits the lungs of patients with cystic fibrosis, is an example. "These bacteria only turn virulent when their local density crosses a certain threshold. At lower densities, the host’s immune system would overwhelm it," he explained.

It is well understood that bacteria use certain signaling molecules to sense density. A positive-feedback biochemical network triggers a switch when the signal concentration is high enough. According to classical control theory, there are many possible networks that yield the same density dependent switching behavior. Evolution, however, has selected one particular network in many different bacterial species. Iyengar is interested in understanding the reasons underlying this selection.

A more complex problem is how bacteria determine the colony’s average temperature in order to optimize their metabolism. This is more difficult than it sounds. Each cell perceives only the temperature around it. Many factors, such as nearby water or chemical reactions, create microclimates that vary significantly from the average.

Iyengar speculates that bacterial colonies use a technique called belief propagation to measure spatial averages. Belief propagation is a well known paradigm from statistical physics that describes how a particle adjusts its behavior based on the behavior of its neighbors.

“If it is used by bacteria, there are measurable consequences that logically follow. We are using our models to guide the type of experiments we need to do to quantify these consequences,” he said.

B. Tech., Indian Institute of Technology, 1993; M.S., Stanford, 1995; Ph.D., 1998
Most of the data created in human history was actually generated in the past handful of years. "Every person in the world on average generates and consumes gigabytes of text, video, Internet media, images, and music every year," Associate Professor Tony Jebara said.

Jebara's specialty is developing machine learning programs that sift through massive amounts of data to discover underlying patterns and make accurate predictions. "I work at the intersection between statistics and computer science, applying machine learning tools to massive data sets where the relationship between variables is often not deterministic. Our algorithms must be fast, because computer speeds are not growing as rapidly as the amount of data they must handle."

Computers slice through data that would take humans years to analyze. Yet their capabilities are only as good as the underlying algorithms, the set of rules used to classify and analyze data. Computers, for example, find it hard to identify faces, a task babies master within months.

This is an area where Jebara made his start by building one of the top face recognition algorithms. His approach to face recognition used probability distributions to calculate the likelihood that two images were of the same individual. Jebara also worked on extending the standard Bayesian algorithms to minimize error rather than maximize likelihood.

Most recently, Jebara has been working on matching and graph algorithms, two promising ways of learning from massive datasets, such as those generated by social networks and the Web. Viewing large amounts of data as a graph often provides a faster and powerful way to solve problems such as data labeling and partitioning.

Also, graphs allow algorithms to be implemented very efficiently by such techniques as message passing, which Jebara has worked on extensively. He has built programs that automatically visualize, label, partition and match data in large data sets, ranging from images to social networks.

Similar algorithms also power Sense Networks, a startup Jebara founded in 2006 to analyze data from telecommunications companies. By tapping smartphone calls and GPS data, Jebara's algorithms can classify people by behavior patterns. Users can then query the network to see where people with similar tastes go to eat, drink, or shop. The phone company can use the data to filter recommendations and provide targeted advertising.

It is one more example of machine learning finding patterns in a world awash with data.

B.S., McGill, 1996; M.S., Massachusetts Institute of Technology, 1998; Ph.D., 2002

Finding Patterns in a Complex World

TONY JEBARA
Associate Professor of Computer Science
many colleges videotape classes so students can review lectures, notes, equations, pictures, presentations, computer screens, and simulations. Yet students rarely use videos to review for exams. Why? Because it takes too long to find the topics and references they need.

John Kender hopes to solve that problem with software that automatically indexes videos. Just like in a book, his index enables people to find exactly what they want in a video. “It is hard to index a video. Most presenters move around and change the subject. Those taping them often lack training. There are none of the classical clues, like fades or establishing shots, to indicate a change in topic. We are developing software to find those clues and create an index.”

Kender has tested the software on videotaped lectures at Columbia. “We have shown that our tools helped students effectively locate the parts of lectures they wanted to study. After we gave them the tools, their grades improved between midterm and final exams. If you have a good way of reviewing educational videos, it pays off in your grades,” he said.

Indexing videos is no simple task. “A professor may start a lecture with a slide, move to a website, then stop to answer a question. That may trigger something he or she forgot to say earlier. They may start four new ideas without finishing previous ones,” Kender related.

Kender’s team found several ways to keep track of this convoluted discourse. One student developed software that recognized and indexed programming languages when they flashed across the screen. “Students can ask for all examples of software code and quickly page through them to find what they want,” Kender said.

Another program reads words from presentations, handwriting from white boards, and captures spoken words using speech-to-text software. Although handwriting and speech-to-text identification is not highly accurate, speakers repeat key concepts often enough to locate them in the video. Another program matches the resulting index with textbook chapters and articles that cover the same material.

A current project involves gestures. Instructors typically use different gestures when reviewing old material, introducing concepts, or working through difficult problems. These gestures vary from teacher to teacher. Kender hopes to decode their meaning by correlating them with such actions as showing new slides, writing on a board, or introducing new words.

Visual indexing could change how people use videos. “It’s like providing a table of contents and index for a book that didn’t have them before,” Kender concluded.

B.S., Detroit, 1970; M.S., Michigan, 1972; Ph.D., Carnegie Mellon, 1980
Securing the Lock after the Key is Stolen

TAL MALKIN
Associate Professor of Computer Science

From online transactions and ATM machines to databases and voting, cryptography lets us share critical information while keeping it safe. Yet cryptographic systems have a weakness. They rely on keys to code and decode messages, and keys can be cracked or stolen.

"Traditional cryptography depends on the assumption that an attacker has no access to secret keys," Associate Professor Tal Malkin related. "Yet sometimes an attacker can hack into a computer or tamper with your hardware. Part of my work is to maintain security even against such adversaries."

She envisions systems that respond when attacked. "We can build systems where the key evolves to protect against an adversary who reads or changes part of the key. Even if an adversary reads the entire key, we can protect future transactions," she said.

Another assumption underlying cryptographic systems is that there is some hard problem that no attacker can solve. For example, the public key software used for secure Internet transactions often relies on the assumption that it is hard to factor the product of two very large prime numbers.

"No one can prove the factoring problem is hard to solve. We assume it is because people have worked on this problem for decades. They have developed sophisticated techniques that are much better than the more obvious approaches, but even those procedures require as many operations as the number of atoms in the universe. But if someone does find an efficient solution, it would break all encryption on the Internet," Malkin said.

As part of her research, Malkin also studies the mathematical foundations of cryptography, searching for the minimal assumptions needed to guarantee security. This starts with studying primitives, such as one-way functions, that act as cryptographic building blocks. "Primitives are small, simple to describe, easy to compute, and hard to crack," Malkin said. "Cryptographers can combine small primitives to form complex, multilayered security systems."

Malkin has also focused on general systems for secure computation among two or more parties, as well as optimizing their performance for specific purposes. One example is the no-fly list. The government wants to keep it secret, while airlines want to protect passenger privacy. Malkin has developed a fast way to exchange critical information without showing compromising data. Other applications of secure computation include online voting, sharing national intelligence, and bidding on projects.

In today's increasingly interconnected world, Malkin's work on provably secure cryptographic protocols could help protect some very important secrets.

B.S., Bar-Ilan University (Israel), 1993; M.S., Weizmann Institute of Science (Israel), 1995; Ph.D., Massachusetts Institute of Technology, 2000
At first glance, the nine-year-old Columbia Newsblaster Web site (newsblaster.cs.columbia.edu) looks like Google News. Both feature the day’s top stories plus sections on national, world, financial, and science/technology news.

The difference is their technologies. Google lists the first sentences of one news article and links to similar stories. Newsblaster publishes summaries of a dozen or more articles – all written and edited by software developed by Professor Kathleen McKeown. “Newsblaster summarizes multiple news articles. We’re using similar technology to answer questions from information on the Web. Today, users read the documents their search returns to see if they are relevant. Our software takes the next step. It looks into the documents, pulls out the relevant information, and summarizes it in a paragraph.” McKeown’s software starts by scraping 25 different Web sites for news every night. It uses key words to cluster articles and categorize topics, counting the number of articles in each cluster to determine its importance.

Once classified, the software uses several approaches to generate summaries. First, it extracts sentences from important sources, such as stories from prominent newspapers and wire services.

It also pairs each sentence with every other sentence in the cluster. It analyzes their similarity and groups related themes together. “The software lines up the sentences in each group side by side and looks at where they overlap or intersect,” McKeown explained. “It is looking for phrases that say the same thing, where words overlap or there is paraphrasing.

“The software parses the sentences for grammatical structure, so it knows that this phrase functioned as a noun and that phrase acts as an adjective. This helps it align similar sentences and fine phrases to create summary sentences. It then generates the summary by ordering the sentences, using information about chronological order of the events. It also edits for coherence, substitutes proper nouns for pronouns, and adds or removes references, depending on whether a person or place is well known or not,” she said.

The core technology has found other uses. A small company is using it to power smart phone applications that track and create timelines for breaking news on specific topics. Another application responds to open-ended questions, generating summaries of information about, for example, a particular event or a particular person. A third creates English summaries from news sources in other languages.

While some Newsblaster stories read like newspaper articles, others are choppier. Still, the technology could become an important tool for making sense of all the information on the Web.

Peer-to-peer networks exploded onto the scene around 2000. That is when Napster, LimeWire, BitTorrent, and similar services made it possible for anyone to download libraries of music and movies for free over the Internet. The new technology gutted music industry profits and led to massive layoffs and downsizing.

Yet peer-to-peer networks are not inherently bad for profits, Associate Professor Vishal Misra argued. In fact, they may prove the most efficient and least expensive way to share media over the Internet.

Back in 2000, large peer-to-peer networks were something new. Instead of warehousing information on a central computer, they took advantage of files distributed on PCs throughout the network to store and send files to other users.

“Smartphones need lots of bandwidth. Wireless providers want us to buy femtocells, small broadcast towers to improve performance in our homes and offices. Instead, they should give us femtocells. Then they could offload traffic from their cell towers and reduce the number of new towers they need to build to support their smartphones,” he said.

“Everyone agrees P2P is a great technical solution. The more users, the more resources the network provides and the faster it responds to requests,” Misra said. This is the opposite of today’s centralized client-server model, which must keep investing in more servers as network demand grows larger. “There shouldn’t be this war between P2P users and people who own music and movie copyrights on the other. We need an economic reboot so that system works for both camps.”

To understand how that might be possible, Misra used game theory to analyze the problem. Ordinarily, models that involve cooperative interactions are extremely hard to calculate, especially for millions of users. Misra simplified those calculations by applying theories based on fluid flow to the continuum of users and peers. “It’s like analyzing a glass of water as a fluid instead of trillions of water molecules. By representing millions of peers as a fluid, it is easier to see their behavior and compute the right incentives.”

For example, Misra estimates that providers of such content as live TV and video-on-demand could save over 90 percent of their Internet distribution costs through user-based P2P networks. “The stores could save lots of money, and people who own legal copies of media might be willing to share them if they receive part of those savings,” said Misra.

“Peers, or users, can help providers reduce costs, as long as incentive structure are in place to reward them,” he concluded.

B.S., Indian Institute of Technology, 1992; M.S., Massachusetts (Amherst), 1996; Ph.D., 2000
Professor Shree K. Nayar’s work is all about seeing things differently. “The basic principles of photography have remained unchanged since the earliest camera obscura,” Nayar explained. “Cameras use an aperture to capture light, a lens to focus it, and some medium to capture the familiar linear perspective image. In the 1990’s, I started asking whether we could use new optics and a computational processing to produce new types of images.”

One of Nayar’s first inventions was the Omnicam. Its combination of lenses and mirrors captures panoramic 360 degree images in a single click. “The image is distorted, since you can’t map a sphere to a flat surface without distortion, but we corrected that with mapping software. In fact, a single 360 degree image could be used to generate any number of traditional views of the scene.

“Placed in the middle of a table of people, it gives the illusion of multiple cameras pointed at individuals during a video conference, although it is one camera with no moving parts,” Nayar said. The camera is also used for surveillance.

Nayar’s next invention, a high dynamic range camera, takes better photographs of scenes that mix dark and light areas. “Let’s say you try to take a picture of a scene with shadows and a bright sky. Today’s digital cameras cannot reveal details within the shadows and the sky. If the sky comes out well, the shadows do not, and vice-versa,” Nayar said.

Nayar’s solution is to use an image sensor with a patterned optical mask on it. The mask ensures that neighboring pixels on the sensor have different sensitivities to light. His software decodes the captured photo to produce one that captures the shaded clouds in the sky and the objects in the shadows. Sony has prototyped the technology for use in its digital cameras.

A third camera enables photographers focus on close-up details without blurring background features. Nayar does this by physically sweeping the image sensor of the camera through an entire focal range, during the exposure of a single photo. The captured photo is again processed by software to obtain one where everything appears in focus.

Nayar has also launched a project to help children around the world learn science, art and culture by assembling and using a digital camera. His BigShot Camera has panoramic and stereo imaging capabilities, and makes it easy to post photos on the Web.

“Each picture is a window on another culture, and youngsters can learn about those cultures from their peers,” he said. To Nayar, it is just another way of seeing things differently.

B.S., Birla Institute of Technology (India), 1984; M.S., North Carolina State, 1986; Ph.D., Carnegie-Mellon University, 1990
Search Google and within a fraction of a second it will return a list of the most popular websites on the subject. Or will it? Assistant Professor Mariana Olvera-Cravioto has been trying to answer that question by understanding what makes a website popular on Google.

The principles behind Google’s page ranking system are well known. It weighs links to and from a page, as well as links of other pages on the same website. Yet the details remain unclear. For example, what counts more, a few links from such important websites as Wikipedia or Technorati, or many links from less significant pages? And what happens to the rankings if you change Google’s search algorithm ever so slightly?

To probe those questions, Olvera-Cravioto relies on a form of probabilistic theory called heavy tail theory. To understand it, consider a normal bell curve. Most samples are grouped close to the center, or mean, and decline rapidly towards the ends of the curve. Heavy tails have far more outliers at the ends of their curves than normal distributions. They are surprisingly common. They show up in the distribution of wealth (few people own more assets), oil reserves (a few have the most value), insurance payouts, and the time supercomputers spend completing tasks.

“Internet video transmission is an example of a heavy-tail distribution. Streaming video only transmits pixels that change. Most of the time, that’s a relatively few pixels. But then the camera changes angles and the whole screen is refreshed. It happens less often, but accounts for most of the transmitted data,” she said. Google search results also have a heavy tail distribution.

“The mathematical techniques needed to solve heavy tails are completely different from what we use with well-behaved distributions,” Olvera-Cravioto said. Those techniques provided some deep insights into Google’s page rankings.

“Before we started our analysis, it was not obvious what determined the relative rankings of websites,” Olvera-Cravioto said. Heavy tail analysis, for example, shows how large numbers of links outweigh important links in popularity rankings.

“Once you understand how it works, you can engineer search algorithms for specific purposes,” she said. “Maybe you want to rank stores by number of sales rather than links, or measure the importance of a paper by how many times it is cited by reliable websites. Adding these things to a search algorithm could make it easier to find the page you’re after.”

B.S., Instituto Tecnológico Autónomo de México, 2000; M.S., Stanford, 2004; Ph.D., 2006

Searching for a Heavy Tail

MARIANA OLVERA-CRAVIOTO

Assistant Professor of Industrial Engineering and Operations Research
Creating the next generation of electronic devices – be they computers, smart phones or displays – will depend on understanding the properties of materials on the nanoscale – one-billionth of a meter.

Professor Aron Pinczuk’s research projects employ advanced optics methods in condensed-matter science, with a focus on understanding the properties of novel materials and the physics of exotic states of matter that emerge in semiconductors at extremely low temperatures. His research findings address issues used by scientists seeking the development of quantum computing and cryptography. The research on graphene, a single atomic layer of graphite, contributes to the quest to initiate a new era in the creation of electronic components.

Pinczuk conducts his research at the Nanoscale Science and Engineering Center at Columbia, in the Department of Applied Physics and Applied Mathematics, and in the Department of Physics. His laboratory had support from the Keck Foundation and his research is funded through the National Science Foundation, the Department of Energy, and the U.S. Office of Naval Research, which support projects that span disciplines in science and engineering.

His research has explored the properties of gallium arsenide, a semiconductor, which is used in advanced optoelectronics, lasers, microwave circuits, and solar cells. To determine material properties in condensed matter systems, he subjects gallium arsenide, to temperatures below 0.1 Kelvin, a temperature at which almost everything freezes. At these temperatures, the electrons cool down to make a liquid, emit light, and exhibit new, unexpected behaviors.

His research with gallium arsenide has also added to the basic science needed to develop a quantum computer, in which computational operations are executed in quantum bits. Theoretical studies show that quantum computers can solve certain problems quicker than classic, digital computer systems.

His findings have also assisted those looking to develop ways to use complex quantum states to build a key used to encrypt computer information. Such encrypted keys could be used to improve the security of computer systems.

Pinczuk’s research with the carbon material, graphene, is part of the effort to develop a new generation of electronics that use carbon components. He studies the properties of carriers of an electric charge as it travels through a single layer of graphene, which is two-dimensional. Scientists are working on larger scale integration of these layers, which will create multi-layer structures with new properties.

“In the case of graphene, there are new properties that develop when you put all the layers together,” says Pinczuk. “It’s a field that is rapidly evolving.”

Licenciado, Buenos Aires (Argentina), 1962; Ph.D., Pennsylvania, 1969
The volume of data we want to analyze is growing even faster than computing power. Professor Kenneth Ross is looking for ways to close the gap. “People are coming up with ever-more challenging database projects, like analyzing the differences in genomes, which have billions of base pairs, among thousands of patients,” Ross said.

Until now, computer scientists have relied on raw increases in computer power to crunch more data. Today, those advances have been harder to achieve. To keep moving forward, engineers reinvented the microprocessor, dividing it into two or more smaller processors, or cores.

Dividing tasks among cores works best when the answers do not depend on the previous step. Databases are like that. “The work you do on one record is pretty much what you do on another, you can process them in parallel,” Ross said.

Yet parallelism comes with its own set of problems, such as cache misses and contention.

Cache misses occur because computer processors have fast and slow memory. They waste hundreds of processing cycles retrieving data from slower memory. Those lost cycles—cache misses—waste half the time needed to perform some tasks.

Ross wants to reorganize data to take up less space in memory. The hard part, he says, is doing this without spending too much time or resources.

“I’m trying to take advantage of relatively recent changes in computer architecture to make database software more efficient. Computer processors are now made up of four to eight smaller processors, or cores. We have to take advantage of those cores by developing code that runs in parallel.”

Contention occurs when several parallel jobs all need to update a single item. “Each of those jobs needs exclusive access to the item for a short time to keep them from interfering with one another. If the item is sufficiently popular, those jobs get stuck in line waiting for their turn to access the data rather than working in parallel,” Ross explained.

Ross’s recent research seeks to automatically detect contention and then create several clones of the busiest data items. “We want to distribute processes among the clones and then combine results. Again, the key is to do this without using more computer resources than we are saving by eliminating contention,” he said.

From genomics to climate, the sciences are accumulating data at a faster rate than ever before. Ross’ work will help make it possible to analyze that data and see what they really mean.

B.Sc., Melbourne, 1986; Ph.D., Stanford, 1991
Imagine a world where library books tell you they are on the wrong shelf and fruit reports it has gone bad to grocers. It is a universe where you can always find your keys or remote control.

This world is under construction in Associate Professor Dan Rubenstein's laboratories. His team is working with small tag-like devices that attach securely to everything from books to baseball bats. "They will let you track all the things you want to track without being tracked by entities you don't want to track you," Rubenstein said.

The devices are called EnHANTS, which stands for energy-harvesting active network tags. "They're designed to soak up energy from the environment to form a network with the tags around them. The networked tags then keep track of one another." Unlike similar radio frequency identification (RFID) tags, which turn on only when activated by powerful radio transmitters, EnHANTS would generate their own power by harvesting energy from ambient light, tiny vibrations, or temperature changes.

Unfortunately, this is not enough power to stay turned on all the time, communicate more than 10 feet, or send lots of information at a time. To get around those limitations, EnHANTS must network with other nearby EnHANTS and devices.

"Existing network protocols waste too much power to work with devices of EnHANTS' size. We have to be more efficient," Rubenstein said. He imagines a room with 10 tagged possessions. The devices sleep to conserve energy, but turn on periodically to see what devices are nearby. Over time, the EnHANTS identify the other devices in the room.

A more powerful device, such as a home wireless network or smartphone, would query the EnHANTS and ask them what they see. Over a period of time, the network would build a map of the room's contents and any sensor data the EnHANTS had to communicate.

"If you start to leave your house and your wallet knows it should be with your belt, coat, and keys, it could tell the network to text a reminder to your cell phone," Rubenstein said.

Meanwhile, Rubenstein's group continues to work on shrinking prototypes to postage-stamp size. "We are really scaling back the components that go into a tag to see how small we can make it," he said. If he succeeds, we may never forget our wallet, keys, or bank cards again.

B.S., Massachusetts Institute of Technology, 1992; M.A., California (Los Angeles), 1994; Ph.D., Massachusetts (Amherst), 2000

Networking Your Wallet, Credit Cards, and Keys

DAN RUBENSTEIN

Associate Professor of Computer Science
The latest smartphones automatically plot your location and update traffic and weather. If a Facebook friend calls, they automatically find and display his or her picture. Slowly, we are weaving together the different strands of the virtual world. Professor Henning Schulzrinne wants to make that fabric richer by making it easier to connect those services and adding sensors to the mix.

Sensors let computers measure and interact with the physical world. “Imagine you’re driving home,” Schulzrinne suggested. “If the temperature is above 80 degrees F, your GPS-enabled cellphone could turn on your air conditioner. It would then turn it off when the last family member leaves home.”

Working behind the scenes, Internet-enabled automation could use sensor data to tailor its response to the situation. Interconnected sensors could warn when household appliances need repairs, water the lawn only when it is dry, analyze traffic so you leave home with enough time to make your dinner reservation, and even check for signs of disease.

“Today, many of these Web services are available to other applications,” Schulzrinne continued. A savvy developer could query a calendar program for today’s appointments or a weather program for a forecast.

“We want to leverage these services into more interesting and comprehensive systems,” Schulzrinne said. “We want to program anything that can be controlled through the Internet, from your lighting and heating to your e-mail and smartphone. We want to make it easier to build smart offices and homes, and to link your calendar with your phone.”

To make that happen, Schulzrinne is focusing on two first steps. One is to develop simple ways to interconnect services, sensors, and applications. “Today, you have to learn Java or other programming languages, or rely on tools from Internet companies. We want to make it easy for the nontechnical to moderately technical users to link things together in interesting ways,” he said.

He is also pushing for standardized interfaces that make it easy to plug sensors into the Web. “There is no reason why every sensor maker should not use the same format to convey information,” Schulzrinne said. “We want to develop a standardized interface, a platform that other people can create modules that use sensor and Internet data to trigger events like services. For example, a module might trigger a stock sale depending on its performance. Another might see if it is going to rain before watering the lawn.”

Ultimately, it could lead to a physical world as interactive as the virtual world that ties it together.

B.S., Technical University of Darmstadt (Germany), 1984; M.S., Cincinnati, 1987; Ph.D., Massachusetts (Amherst), 1992
In “20 Questions,” one player thinks of an object and the others get 20 yes-no questions to guess its identity. “That’s easy, but what if you let the answerer lie three times? That makes it much more difficult,” Associate Professor Rocco Servedio said.

That is the type of problem researchers face when false signals, or noise, corrupt data. Servedio’s goal is to develop robust algorithms that learn complicated rules even in the presence of noisy data. Such algorithms could learn patterns that improve sensor performance, predict earthquakes, or forecast financial markets.

One of Servedio’s most powerful tools is geometry. “When you cast a learning problem in a geometric framework, you’re often on the way towards solving it,” he said.

Imagine, for example, a piece of paper with red plus signs and green minus signs on opposite sides of an unknown dividing line. A few pluses are mixed with the minuses and vice versa. “In this two-dimensional example, you can eyeball the data and see which points don’t belong. In higher dimensions, where each point has many coordinates, this is much more difficult, though we can sometimes pull it off with tools from high-dimensional geometry,” he said.

“The way people understand something is by drawing pictures. I’m usually working in high-dimensional Euclidean spaces where it’s tough to draw accurate pictures, but thinking geometrically still provides useful insights.” Servedio also takes a geometric approach to studying rules used to classify information. One popular approach is the decision tree. Like 20 Questions, it uses a sequence of yes-no questions to decide how to label data points.

“If you think of this logical representation geometrically, you can sometimes see properties that would have otherwise remained hidden. These insights can lead to better learning algorithms,” Servedio said.

Servedio also uses geometry to compensate for missing data. Imagine that it takes 1,000 coordinates to describe a data point completely. What kind of learning is possible if only one of those coordinates is available?

“There are ways to compensate for massive amounts of missing data,” Servedio said. “It might sound impossible, but doctors do something like this all the time. They could potentially run thousands of clinical tests on a patient to fully describe his or her condition, but a good doctor can make a useful diagnosis from just one or two tests.”

A.B., Harvard, 1993; M.S., 1997; Ph.D., 2001
Professor Cliff Stein has built a career on finding algorithms to solve difficult problems—but not precisely. Stein specializes in algorithms that estimate the answer to problems that are difficult to solve. In operations research and computer science, these are problems that grow exponentially more complex as the number of inputs grows.

This contrasts with simple problems, like alphabetizing words. Double the number of inputs—words—and it takes only about twice as long to accomplish the task.

A well-known difficult problem is calculating the most efficient route for a salesman to visit different cities. To find the most efficient route, a computer must calculate all possible outcomes. For five cities, there are 120 potential paths. For 10 cities, 3.6 million. “For 80 cities, there are roughly as many possible answers as there are atoms in the universe,” Stein said.

“No conceivable advance in computing power would enable us to solve that problem precisely,” Stein noted. “But if you’re willing to solve it approximately, you can do so more easily and efficiently.”

Many algorithms already exist for estimating the solution for the traveling salesman and other difficult problems. Stein prefers to break new ground, studying the fundamental structure of problems to develop new algorithms.

“There’s a collection of algorithmic tools that are commonly used to solve many problems. But often there are problems that are important to solve. It is worth investing the time to study their mathematical or combinatorial structure to come up with a solution specific to that problem.”

Much of his work deals with scheduling everything from computer systems to factories. Scheduling starts with jobs and the machines needed to complete them. Constraints—jobs take different amounts of time, some are more important than others, some tasks depend on others—add to the difficulty. So do different objectives, like fast completion, minimal resources, and rapid response.

Stein is looking at ways to apply scheduling to computer processors in order to save energy. “Most chips can run at four or five different speeds. If you run at half speed, you decrease energy use by roughly a factor of four. But no one has figured out how to give chips the intelligence to know when to slow down, so they typically run at top speed all the time,” he said.

By estimating a chip’s workload, constraints, and performance goals, Stein believes he can achieve significant energy savings. Even if his estimates are not precise.

B.S.E., Princeton, 1987; M.S., Massachusetts Institute of Technology, 1989; Ph.D., 1992
Professor Joseph Traub is best known as a pioneer in the computational complexity of continuous problems. This involves understanding the least amount of resources—time, memory, communications—needed to solve a computing problem. “My strategy is to start a new area of research or get into something fairly early. Then I can just walk along and pick up diamonds of knowledge and insight. I never have to strip mine for them,” he said.

It is probably as good a background as any for his investigations into the potential of quantum computing. It is a quest at the intersection of physics, mathematics and computer science.

Quantum computing stands conventional computing on its head. For example, bits are the basic unit of information in today’s computers. They can have one of two values, either zero or one, which microprocessor transistors represent as on or off.

Quantum computers are built around qubits, which have a property called superposition. This means they can be in many quantum states between zero and one, all at the same time. The more qubits a processor has, the more potential states it allows.

Qubits also have a property called entanglement. For reasons not yet understood, changing the quantum state in one of two entangled particles instantaneously changes it in the other. “That enables qubits to work together without wires,” Traub said.

Because quantum computers are not limited to on-off states, they can calculate many possible answers at once. This could make it possible to calculate very complex problems rapidly.

“What I’m trying to do is ask, ‘Where are the big wins?’ In particular, what kind of problems could a quantum computer solve that physicists and chemists are really interested in solving,” Traub said.

One of those problems is calculating the lowest energy state, or ground state energy, of a large number of particles. “This is a central problem in computational chemistry, and it would allow us to predict chemical reactions better,” Traub explained.

The problem, he explained, is that ground state energy calculations are difficult and soak up computer resources. A quantum computer’s ability to make multiple calculations simultaneously could give chemists the tool they need to predict particle interactions in large systems.

“We’re theoreticians, trying to understand the type of problems quantum computers might be able to solve. Physicists may never succeed in building one, but if they do, we want to be ready,” Traub concluded.

B.S., College of the City of New York, 1954; M.S., Columbia, 1955; Ph.D., 1959
“When the solution is simple, God is answering,” Einstein once commented. He believed we could discover nature’s laws only when they connected a few variables, like the relationship between temperature and pressure or energy and mass. “When the number of factors coming into play is too large, scientific methods in most cases fail,” Einstein said.

Of course, Einstein did not have computers. Professor Vladimir Vapnik does. He works in machine learning, a discipline that uses algorithms to detect automatically those laws of nature that depend on hundreds or even thousands of parameters. This enables computers to make better predictions, and also provides insights into the elusive nature of human learning.

Today’s machine learning technology requires many examples to generate accurate rules. Yet humans clearly learn to understand their complex world from far fewer examples. This led Vapnik to consider how teachers provide students with what he calls “privileged information,” holistic knowledge often delivered as metaphors and comparisons.

Master classes for musicians are an example. “The teachers cannot show students how to play an instrument because their technique is not as good. Instead, teachers may use metaphors or comparisons to show students how to understand a piece. This may sound like nonsense in terms of musical technique, but it helps them play better,” Vapnik said.

Vapnik has shown mathematically that privileged information could slash the samples needed for machine learning by the square root of the original number. “Instead of ten thousand examples, we would need only one hundred,” he said.

He demonstrated this using privileged information to help a computer identify handwritten numbers. He asked Professor of Russian Poetry Natalia Pavlovitch to write a short verse describing her feelings about each number sample. The information was subjective and not available by analyzing only the numbers. Including it during training yielded more accurate results than training with the numbers alone.

Vapnik also used surgeons’ descriptions of biopsy pictures—from “quiet” to “wide aggressive proliferation”—to improve the classification of tumors. The notes were impressionistic, but improved the computer’s ability to identify cancerous cells.

Humans frequently use such holistic privileged information to make sense of complex phenomena. Providing it to machines could open a new door onto a complex universe.

“For 2,000 years, we believed logic was the only instrument for solving intellectual problems. Now, our analysis of machine learning is showing us that to address truly complex problems, we need images, poetry, and metaphors as well,” Vapnik concluded.

M.S., Uzbek State University, Samarkand, 1958; Ph.D., Institute of Control Sciences, Moscow, 1964
A dvancing wireless communication technology to a new generation of application and service is one of today's prime research disciplines. Demands for higher capacity drive the need to create novel signal transmission techniques and advanced receiver signal processing methods. Challenging design requirements are compounded by the complexity of the nature of the transmitter and receiver: a complicated system consisting of radio frequency, analog and mixed-signal components. Plus, heated competition in the development arena forces tight time-to-market deliverables. To develop effective next-generation wireless technology under the constraint of thousands of variables, it is important to use mathematical modeling and analysis, computer simulations, fast calculations and data summaries to thoroughly account for manufacturing process variations before build-out. Using these tools to analyze production provides a comprehensive transistor-level statistical design and verification framework. With it, designers can troubleshoot and devise design enhancements to solve the issues of fading, impulsive noise, and co-channel interference in the concept phase.

Dr. Xiaodong Wang is a leading researcher in signal processing, computing and communications. His broader research interests include information theory, algebraic coding theory, wireless communications, optical communications, communication networks, statistical signal processing, and genomic signal processing. Results of his research have included extensive publication in these areas, most recently in the areas of chip-level asynchronism on a Code Division Multiple Access (CDMA)-based overlay system for optical network management; modulation classification via Kolmogorov-Smirnov test; Generalized Likelihood Ratio Test (GLRT)-based spectrum sensing for cognitive radio with prior information; and blind frequency-dependent I/Q imbalance compensation for direct-conversion receivers.

Professor Wang also has become active in the emerging field of genomic signal processing (GSP). The aim of GSP is to integrate the theory and methods of signal processing with the global understanding of functional genomics, with special emphasis on genomic regulation. He took part in an NSF-funded multidisciplinary collaborative project to develop a structural health monitoring (SHM) system using a wireless piezoelectric sensor network.

Dr. Wang is a Fellow of the Institute of Electrical and Electronic Engineering (IEEE). He received the 1999 NSF CAREER Award and the 2001 IEEE Communications Society and Information Theory Society Joint Paper Award. He has served as an Associate Editor for the IEEE Transactions on Signal Processing, the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, and IEEE Transactions on Information Theory. He is listed as an ISI-Highly-Cited researcher.

B.S., Shanghai Jiaotong University, 1992; M.S., Purdue, 1995; Ph.D., Princeton, 1998

Devising a Design Framework for Next-Generation Wireless Technology

XIAODONG WANG
Professor of Electrical Engineering
Predicting Waves Mathematically

MICHAEL WEINSTEIN
Professor of Applied Mathematics

Understanding the behavior of waves in complex environments holds the key to advances in a wide range of applications—from optical communications and computer technology to the prediction and detection of seismic, atmospheric, and oceanic phenomena. Wave phenomena are described using partial differential equations, which are a mathematical encoding of physical laws.

But significant challenges arise because phenomena are both multiscale—they derive from activity and interactions among very small spatial scales all the way up to very large scales—and nonlinear, which leads to waves that distort dramatically and scatter differently as their size is changed. These general features limit the solvability of problems on even the fastest computers.

Professor Michael Weinstein develops hybrid analytical/computational approaches, which combine asymptotic mathematical analysis with computer simulation. Asymptotic analysis yields approximate, but fairly explicit and detailed information, on the very small-scale phenomena. With these degrees of freedom “solved for,” the computer can then focus on the larger scales and efficiently give approximate, yet very accurate predictions.

Applying these approaches to the partial differential equations of optics, Weinstein has discerned how “soliton” light-pulses travel and interact within communications lines. This work has a wide range of practical applications: from determining the stability of optical pulses to ideas on how to robustly encode information in streams of optical pulses.

He has proposed designs of novel optical media to slow or even stop light pulses in micro-structured waveguides, and has proposed their application to optical buffering of information. A recent project exploits parallels between the equations of electromagnetics with those arising in the theory of shock waves in supersonic flight, to understand the generation of broadband, multi-colored light from laser light of a single color. Broadband light sources have applications ranging from communications to imaging science.

In other recent work, he is addressing concerns metamaterials: specially engineered microstructures, which act as a macroscopic device, and achieve properties not possible using naturally occurring materials. One application studied by Weinstein is the attainability of the cloaking effect. Cloaking involves surrounding a region of space by an appropriate metamaterial. Anything in the surrounded region is undetectable by exterior sensors, and anything within the shielded region is isolated from the exterior world.

Other application areas of metamaterials envisioned include improved solar energy cells, secure communications and sensors.

Weinstein is a Fellow of the Society for Industrial and Applied Mathematics (SIAM), elected for his “contributions to the analysis and applications of nonlinear waves.”

B.S., Union College, 1977; M.S., Courant Institute-NYU, 1979; Ph.D., 1982
Computers are solving ever more complex problems, yet some problems have resisted intense efforts for many decades. How can we tell which problems can be solved efficiently and which cannot? How do we find the most efficient algorithms? And for intractable problems, how do we find the best solutions possible in reasonable amounts of time? These are some of the challenges taken on by Professor Mihalis Yannakakis.

One line of his research seeks to understand the inherent computational complexity of problems. “It turns out that many computational problems from diverse fields are intimately related to one another,” he said. For example, optimizing network designs, scheduling jobs, and folding proteins all exhibit essentially the same type of computational difficulties. Yannakakis seeks to find the underlying features that characterize the complexity of different problems and identify their unifying principles.

Many optimization problems are computationally hard, in the sense that we cannot compute efficiently the optimal solution. For these cases, Yannakakis has been working on algorithms that compute near-optimal solutions. His goal is to design efficient approximation algorithms with provable performance guarantees.

Yannakakis’ third research thrust involves tradeoffs when making decisions. “We care about a design’s quality and also its cost, or a health treatment’s benefits and risks. Typically, there is no one solution that is optimal for all criteria, but rather many incomparable solutions that encapsulate the tradeoffs between different criteria,” he explained.

For two criteria, Yannakakis visualizes these tradeoffs as a curve on the plane. As the number of criteria rise, the tradeoffs form a surface in a higher dimensional mathematical space. “It is generally impossible to generate all the points on the tradeoff surface because there are usually an exponential or even infinite number of them. But we wish to generate enough points to represent the whole design space, so decision makers have an accurate enough view of the tradeoffs to make an informed choice,” he said.

Yannakakis’ approach is to design algorithms with guaranteed succinctness and accuracy to compute a carefully selected small set of solutions that offer the best possible representation of that space.

“Computers could work forever on a task,” he summarized. “We try to characterize what we can actually compute efficiently for a specific task and in general. Like trying to understand the physical world, we’re trying to find the laws that govern the computational world. We want to determine the powers and limitations of computation.”

Dipl., National Technical University of Athens (Greece), 1975; M.S., Ph.D., Princeton, 1979

Calculating What Is Possible
MIHALIS YANNAKAKIS
Percy K. and Vida L. W. Hudson Professor of Computer Science and Professor of Industrial Engineering and Operations Research
The design and deployment of mobile and wireless networks has undergone an extraordinary transformation. While this technology already forms the backbone of crucial systems such as healthcare, disaster recovery, public safety, manufacturing, and city-wide broadband access, it has even greater potential. The flexibility inherent in cellular, sensor, mobile ad hoc, mesh and wireless local area network technologies delivers an almost endless range of applications, including mobile banking, inter-vehicle communication, space exploration, and climate-change tracking.

Despite their promise, efficiently controlling wireless networks is a challenging task, due to interference between simultaneous transmissions, mobility of the nodes, limited capacity of the wireless channel, energy limitations of the devices, and lack of central control. Such distinct characteristics set wireless networks apart from other networking technologies and pose numerous challenging theoretical and practical problems.

To tackle these problems, Assistant Professor Gil Zussman focuses on designing new wireless networking architectures and on improving the performance and resilience of existing networks. Due to the special characteristics of these networks, Zussman designs architectures and algorithms that are optimized across multiple layers of the networking protocol stack. For example, he has been working on energy-aware protocols that take into account energy consumption and battery status while making joint decisions regarding routing and scheduling. Zussman has been recently focusing on developing algorithms and prototypes for Energy Harvesting Active Networked Tags (EnHANTs). These tags harvest their energy from the environment and can be used in various tracking applications, and particularly in disaster recovery applications.

Moreover, in order to enable the efficient operation of distributed algorithms which usually have inferior performance to centralized algorithms, Zussman has been working on identifying topologies in which distributed algorithms obtain maximum throughput. His results in this area enable the partitioning of networks to subnetworks in which distributed algorithms operate very well, thereby improving the overall network performance.

Other research projects of Zussman’s group focus on controlled mobility of wireless nodes, dynamic spectrum allocation and cognitive radio, interfaces between wireless and optical networks, and resilience of networks to geographically correlated failures. Results regarding the latter include identifying vulnerabilities of networks to large scale attacks, such as Electromagnetic Pulse (EMP) attacks, and mechanisms to mitigate the effects of such attacks.

Zussman was a Postdoctoral Associate with the Massachusetts Institute of Technology (MIT) as a Fulbright Fellow and Marie Curie Fellow. He is a senior member of the Institute of Electrical and Electronic-Engineers (IEEE).

B.Sc., B.A., Technion-Israel Institute of Technology, 1995; M.Sc., Tel Aviv University, 1999; Ph.D., Technion-Israel Institute of Technology, 2004