



In the field of
human-computer interaction,
computer science
meets human behavior.

Dancing

<!-- WITH COMPUTERS -->

By Kathleen McGarvey

Go into any public space and look at the people around you. Odds are, some if not most of them will have their neck craned downward, their eyes lowered, and one hand cradling their phone.

You're looking at one of the primary relationships in the lives of many people today. But let's face it: it's not quite rapport. Conversations with our computers are pretty one-sided. Even storied innovations in voice recognition—Hello, Siri?—are often frustrating and fruitless.

"Every other day, I feel like throwing my laptop out the window because it won't do what I want it to do," says Henry Kautz, the Robin and Tim Wentworth Director of the Goergen Institute for Data Science and professor of computer science. He's an expert in artificial intelligence—so if throwing your laptop or smartphone out the window has crossed your mind on occasion, too, well, at least you're in good company.

What if relating to computers were more like the way we communicate with other people?

That's a vision that scientists in the field of human-computer interaction, or HCI, are working to realize. It's an ambitious goal, but they're making significant headway.

Philip Guo, assistant professor of computer science and codirector of the Rochester Human-Computer Interaction Lab, calls HCI a blend of science and engineering.

"It's about attempting to understand how people interact with computers—that's the science part—and creating better ways for them to do so. That's where engineering comes in," he says.

The field emerged around the 1980s, with the rise of personal computing and as the work of computer scientists began to be informed by cognitive science. Anyone who can recall

<!-- It's like a dance. I say something; you understand what I'm trying to say; you ask a follow-up question; I respond to that. But a lot of the things are implicit. And that entire richness of conversation is missing when you interact with a computer. -->

EHSAN HOQUE, ASSISTANT PROFESSOR OF COMPUTER SCIENCE AND ELECTRICAL AND COMPUTER ENGINEERING

the labor of entering DOS commands to complete even the simplest tasks knows well the trajectory computers have taken toward their more intuitive configurations today.

The issues that HCI experts at Rochester are investigating range widely: improving online education, helping people to communicate more effectively, monitoring mental health, and predicting election outcomes.

Personal Communication Assistance

"I like to build interfaces that allow people to interact with computers in a very natural way," says Ehsan Hoque, assistant professor of computer science and electrical and computer engineering.

And what would an "unnatural" way be? That's the way we use computers now, he says.

When we talk with someone, we use not only words, but also facial expressions, patterns of stress and intonation, gestures, and other means to get our points across.

"It's like a dance," says Hoque, who codirects the HCI lab with Guo and Kautz. "I say something; you understand what I'm trying to say; you ask a follow-up question; I respond to that. But a lot of the things are implicit. And that entire richness of conversation is missing when you interact with a computer."

Much of what we communicate, and what others communicate to us, isn't registered by our conscious minds. Eye contact, smiles, pauses—all speak volumes. But most of us have little idea of what we actually look like when we're speaking with someone. Our own social skills can be a bit of a mystery to us.

So Hoque has developed a computerized conversation assistant—called "LISSA" for "Live Interactive Social Skills Assistance"—that senses the speaker's body language and emotions, helping to improve communication skills. The assistant, who looks like a college-age woman, evaluates the nuances of the speaker's self-presentation, providing real-time feedback on gestures, voice modulation, and "weak" language—utterances such as "um" and "ah." Intriguingly, the system allows people to practice social situations in private.

The first iteration of the project was Hoque's doctoral thesis at MIT. There he tested the system—then called MACH, for My Automated Conversation Coach—on MIT undergraduate job seekers. Career counselors found the students who had practiced with MACH to be better job candidates. He has since tested the technology with date-seekers for speed-dating, in a study designed with Ronald Rogge, associate professor of psychology, and Dev Crasta, a psychology graduate student. Their study showed that coaching by LISSA could help online daters subtly improve eye contact, head movement, and other communicative behaviors.

Hoque is also adapting it for use by people with developmental disorders, such as autism, to help them enhance their interactions with others.

People with autism often have an "unusual inflection or intonation in their voice—it's one of the things that interfere with their social communication," says Tristram Smith, a professor in the Department of Pediatrics and a consultant on the project.

Job interviews can be very difficult for people with autism. "We don't have a lot of interventions to help with their conversational

skills, and problems with conversational speech are really at the core of what autism is," he says.

But computers are well suited to assisting. They're better at analyzing speech patterns than people are, and they can show children what happened when they spoke, bringing together as a visual display the words they uttered and the gestures they made.

Helping people become better communicators is a project close to Hoque's heart.

"I have a brother who has Down syndrome," he says. "He's 15, he's nonverbal, and I'm his primary caregiver." When he was doing his doctoral research at MIT, Hoque knew that he wanted to build technology that benefits people in need and their caregivers. He worked on assisted technology to aid people in learning to speak effectively, improve their social skills, and understand facial expressions in context.

From that work, he has created other tools, such as ROCSpeak, which aims to help people become better public speakers by analyzing the words they use, the loudness and pitch of their voice, their body language, and when and how often they smile. He's even developed "smart glasses" that provide speakers with real-time, visual feedback on their performance. That system is called "Rhema," after the Greek word for utterance.

The United States Army has funded Hoque's work to use the technology to study deceptions. "We can say it's out-of-sync behavior, so it could be deception, it could be stress, it could be nervousness," says Hoque. "But when the behavior is getting out of sync—when your speech and your facial expression are not in sync—something is wrong, and we can predict that."

Analyzing Social Networks

Predictions are the core of the work of Hoque's colleague, Jiebo Luo. A professor of computer science, Luo has many projects afoot.

In one of them, he is working—with researchers at Adobe Research—to harness the power of data contained in the sea of online images by training computers to understand the feelings that the images convey. For example, the photos of political candidates that people choose to post or share online often express information about their feelings for the candidate.

By training computers to digest image data, the researchers can then use the posted images to make informed guesses about a candidate's popularity.

A team led by Luo and Kautz is using computers to improve public health. Their "Snap" project uses social media analytics for a variety of health applications ranging from food safety to suicide prevention.

They're also investigating how computers can help in diagnosing depression by turning any computer device with a camera into a tool for personal monitoring of mental health. The system observes the user's behavior while using a computer or smartphone. It doesn't require the person to submit any additional information.

"There's proof that we can actually infer how people feel from outside, if we have enough observations," says Luo.

Through their cameras, devices can look back at us as we view their screens—and extracting information from what the camera



“sees” allows the device to “build a picture of the internal world of a person,” he says.

The camera can measure pupil dilation, how fast users blink, their head movement, and even their pulse. Imperceptibly to the casual observer, skin color on the forehead changes according to blood flow. By monitoring the whole forehead, the computer can track changes in several spots and take an average. “We can get a reliable estimate of heart rate within five beats,” he says.

Online Learning

While Luo’s work, like Hoque’s, turns the computer into an observer of human behavior, Guo’s research uses computers to bring people metaphorically closer together. His focus is online education.

“I’m trying to humanize online learning,” he says.

It’s easy to put videos, textbooks, problem sets, and class lecture notes online, he says, but the simple availability of materials doesn’t translate into people actually learning online. Education research has shown that motivation is a decisive factor—motivation that is effectively instilled through small classes and one-on-one tutoring.

“The challenge of my research is how do you bring that really intimate human interaction to a massive online audience,” Guo says. He’s working to build interfaces and tools that will bring human connection to large-scale online education platforms.

He has already made headway in a website for people learning the popular programming language called Python. His free online educational tool, Online Python Tutor, helps people to see what happens as a computer executes a program’s source code, line by line, so that they can write and visualize code. The site has more than a million users—enormous, for a research site—from 165 countries. Guo is working to connect people through the site, so that they can tutor each other, even though they may be continents apart.

A ‘Grand Challenge’

Hoque’s work with Smith on developing the communication assistant for use by children on the autism spectrum is part of a collaboration with colleague Lenhart Schubert, a professor of computer science, that has received funding from the National Science Foundation. They aim to improve the assistant so that it can understand—at least in some limited way—what a user is saying and respond appropriately.

They began testing the language comprehension part of the system in the speed-dating experiment. There they had a person helping the computer to provide appropriate responses, in what’s known as a “Wizard of Oz” study, in which an operator is controlling the avatar from behind the scenes. Now they’re in the process of automating the system.

But the problem of natural language processing for computers is a thorny one. Teaching computers to understand spoken language has preoccupied artificial intelligence researchers since at least the 1960s.

Computers now can recognize speech within a relatively limited domain. You can ask your smartphone for help in finding a Chinese restaurant or ask it to help you make an airline reservation to fly to Los Angeles. But when it comes to the kind of dialogue that people actually have—not narrowly focused but free-flowing and context-dependent—it’s much harder to predict what’s going to be said, and the computer is operating according to predictions. “Basically that’s been beyond the capacity of artificial intelligence for all these decades,” says Schubert.

Improvements in machines’ ability to parse the structure of language have moved the project forward—but not far enough. In a sentence that’s 20 words long, typically the machine will make a couple of mistakes.

Consider the sentences “John saw the bird with binoculars” and “John saw the bird with yellow feathers.” You know that it’s John, and not the bird, who has the binoculars—and that it’s the bird, not John,

that sports the feathers. And the seemingly simple question “What about you?” calls for very different answers depending on whether the previous sentence was “I’m from New Jersey” or “I like pizza” or “I’m studying economics.” But such context-dependent information is much more elusive for computers than it is for people.

“The system has to have world knowledge, really, to get it right. And knowledge acquisition turns out to be the bottleneck,” Schubert says. “It has stymied researchers since the beginnings of artificial intelligence.” He calls it “the grand challenge.”

But a grand challenge is there even for the nonverbal part of the equation, says Hoque. “Even something simple like a smile: when you smile, it generally means you’re happy—but you could smile because you’re frustrated; you could smile because you’re agreeing with me; you could smile because you’re being polite. There are subtle differences. We don’t know how to deal with that just yet. So there’s still a long way to go on that, too.”

The HCI program graduated its first crop of doctoral students last spring. Erin Brady ’15 (PhD)—whose research is concerned with using technology and social media to support people with disabilities—is now an assistant professor at Indiana University-Purdue University Indianapolis. Yu Zhong ’15 (PhD) works on mobile apps for accessibility and on ubiquitous computing—inserting microprocessors in everyday objects to transmit information. Google Research has hired him as a software engineer.

The third member of the class, Walter Lasecki ’15 (PhD), is now in his first year as an assistant professor of computer science and engineering at the University of Michigan. He began his doctoral work at Rochester in artificial intelligence but moved to HCI to explore how crowds of people, “in tandem with machines, could provide the intelligence needed for applications we ‘wish’ we could build,” he says. “I realized that much of what we know how to do, what we think about how to do, is limited by what we can do using automation alone today,” he says. Combining computers with human effort—in what’s called “human computation”—is essentially letting researchers try out new system capabilities.

“It lets us see farther into the future. We can deploy something that works, something that helps people today. And as artificial intelligence gets better, it can take over more of that process.”

What initially drew Lasecki to HCI was his interest in “creating real systems—systems that have an impact on real people,” he says. Kautz cultivated this focus in the computer science department by hiring first Jeff Bigham—now at Carnegie Mellon—and then Hoque and Guo. As Bigham was, they’re concerned with practical applications. “That’s certainly a strength, this focus on applications and system building,” Lasecki says of Rochester’s program.

Those systems will become an ever-more pervasive part of our lives, Hoque predicts, and the field of HCI will gradually become an essential part of other disciplines. In fact, it’s already happening. More than half of the students in HCI courses at Rochester aren’t computer science majors. They’re from economics, religious studies, biology, business, music, studio arts, English, chemistry, and more.

Hoque quotes the founder of the field of ubiquitous computing, Mark Weiser, who once wrote, “The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.”

Computing, Hoque says, is on its way to becoming like electricity: it’s everywhere, but you don’t really see it.

“We won’t see it, we won’t think about it. It will just be part of our interaction—maybe part of our clothing, part of our furniture. We’ll just interact with it using natural language; it will be natural interaction.

“And we’re working toward that future.” 

ROCSpeak is available for use at <https://www.rocspeak.com>.

