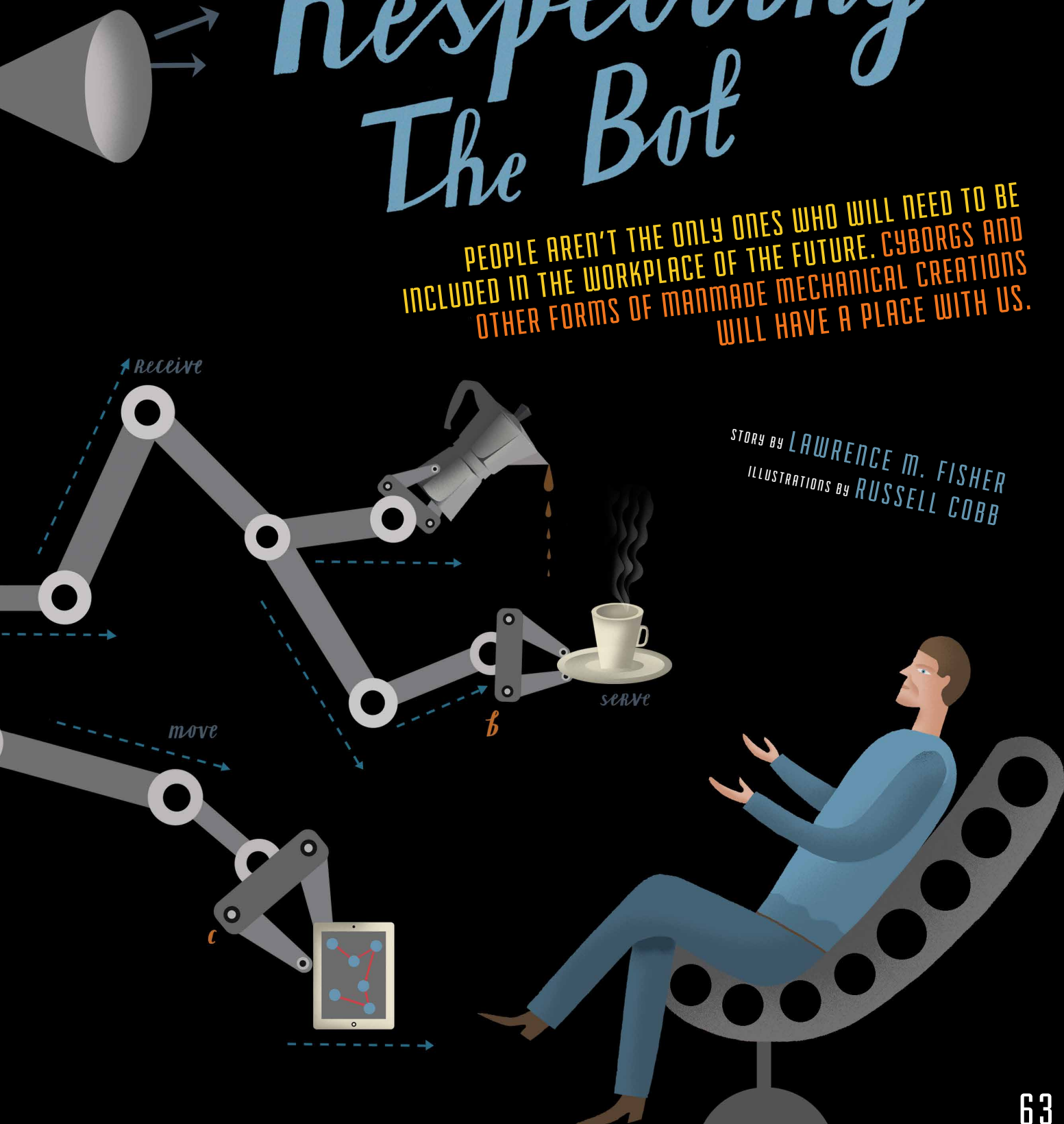# Respecting The Bot

PEOPLE AREN'T THE ONLY ONES WHO WILL NEED TO BE INCLUDED IN THE WORKPLACE OF THE FUTURE. CYBORGS AND OTHER FORMS OF MANMADE MECHANICAL CREATIONS WILL HAVE A PLACE WITH US.

STORY BY LAWRENCE M. FISHER

ILLUSTRATIONS BY RUSSELL COBB
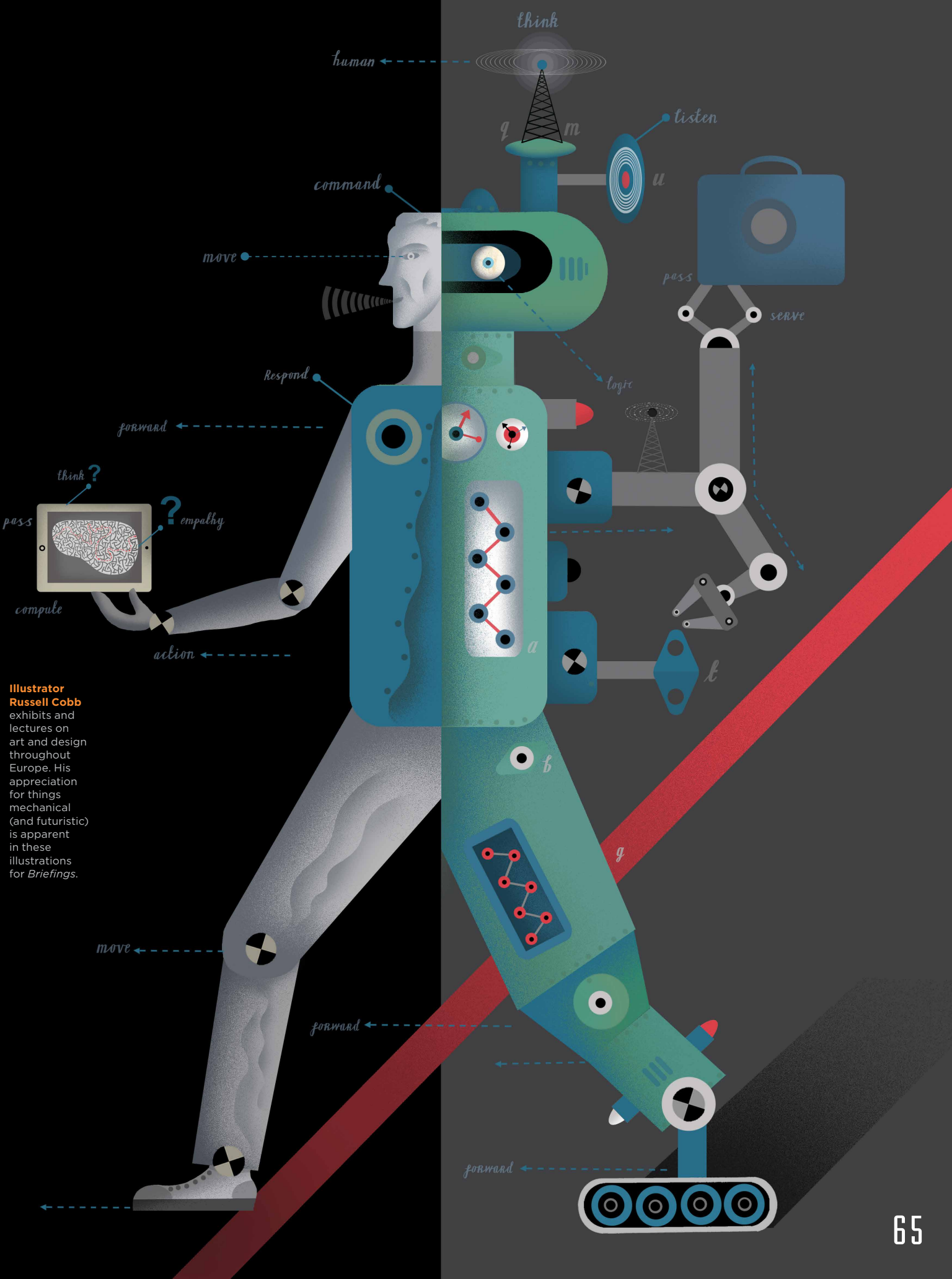
Receive

move

serve

IN THE PROJECT TEAMS SO COMMON IN THE MODERN ENTERPRISE, PARTICIPANTS ENGAGE WITH CO-WORKERS FROM DIVERSE CULTURES, FIND COMMON CAUSE AND PURSUE SHARED GOALS. AS ROBOTS AND ARTIFICIAL INTELLIGENCE SYSTEMS ENTER THE WORKPLACE, EMPLOYEES MAY NEED TO LEARN TO COLLABORATE WITH NON-HUMAN COLLEAGUES.

The common wisdom is that artificial intelligence has overpromised and underdelivered, and it is surely one of the most-hyped technological developments of all time.

Your first robotic colleague will probably not resemble the paranoid HAL of Stanley Kubrick's "2001: A Space Odyssey," or the sultry-voiced Samantha of Spike Jonze's 2013 film, "Her," but it could be something more akin to an artificial administrative assistant. That was the goal of CALO, or the "Cognitive Assistant that Learns and Organizes," an SRI International project funded by the Defense Advanced Research Projects Agency. CALO ran from 2003 to 2008 and produced many spinoffs, most notably the Siri intelligent software assistant now included in Apple iOS.

Siri is not especially human-like, but successors currently in development will be much more so. As the interface with these devices moves from command-driven to conversational, our relationships with them will inevitably become more interactive, even intimate. Even if the devices are not truly sentient, or conscious, research shows that we will experience them as if they were. Firms now provide diversity training for working with different races, gender identities and nationalities. Will we need comparable workplace policies for human-robotic interaction?

think

human ←

listen

command

move ←

q m u

pass

serve

Respond

forward ←

logic

think ?

pass ? empathy

compute

action ←

Illustrator Russell Cobb exhibits and lectures on art and design throughout Europe. His appreciation for things mechanical (and futuristic) is apparent in these illustrations for *Briefings*.

a

b

l

g

move ←

forward ←

forward ←

Any good article about robots should mention Isaac Asimov's three fundamental Rules of Robotics, and this is as good a place as any: One, a robot may not injure a human being, or, through inaction, allow a human being to come to harm. Two, a robot must obey the orders given it by human beings except where such orders would conflict with the First Law. Three, a robot must protect its own existence as long as such protection does not conflict with the First or Second Laws. Asimov first postulated the rules in a 1941 short story, "Runaround," which happened to be set in the year 2015, surely a sign that the time has come to at least think about fundamental rules of human-robotic interaction.

What will it be like to work side by side with robots? Actually many of us already do every day; we're just not always aware of it. When you run the spelling and grammar check on a document you produced with Microsoft Word, an artificial intelligence is doing the job of copy editor. If you drive one of the new Volvos equipped with IntelliSafe, the car will hit the brakes or steer its way out of danger far faster than a human being can respond. You may be holding the wheel, but at critical moments, the robot is driving. Google's self-driving cars take over driving entirely — no need to hold the wheel at all.

And if you use one of those clever scheduling apps that pluck dates, times and phone numbers from your e-mails to set up meetings and articulate goals and objectives, congratulations, you already have an artificial intelligence on your project team. It may lack a face, a voice and a personality, at least for now, but it is already performing important tasks that once required a human co-worker.

All of the systems just described are the products of fairly mainstream artificial intelligence programming. The reason they work now when they didn't work in past decades is primarily due to the explosive growth in processing power available at an ever-lower cost. When your laptop or smartphone has more processing power than the Apollo mission,

amazing things become possible. That doesn't give systems feelings, or give them the power to learn the way the human brain does, but it does make them very capable machines.

While conventional AI has been about programming ever-more powerful computers to tackle ever-more complex problems, an emerging technology called artificial general intelligence, or AGI, is about developing systems with less ingrained knowledge, but with the capacity to learn by observation. You may well need to mind your manners around these machines.

"Whether you're polite to your software will actually make a difference in the future," said Pei Wang, a professor of computer science at Temple University. "But that assumes that the system will actually learn. It will need to have general intelligence, the ability to evaluate someone's performance, some form of emotion."

Even without emotion, which even AGI advocates say is some distance away, learning computers may respond more productively when they are treated better. Even if the robot does not care about your rude behavior, your human co-workers will feel uncomfortable, which is not conducive to team solidarity. Owing to something computer scientists call the ELIZA effect, people tend to unconsciously assume computer behaviors are analogous to human behaviors, even with quite unsophisticated systems, like the ATM machine that says "thank you." Teammates may assume you have hurt your artificial admin's feelings, even if she has none.

ISAAC ASIMOV'S THREE FUNDAMENTAL LAWS OF ROBOTICS

1 A ROBOT MAY NOT INJURE A HUMAN BEING, OR, THROUGH INACTION, ALLOW A HUMAN BEING TO COME TO HARM.

2 A ROBOT MUST OBEY THE ORDERS GIVEN IT BY HUMAN BEINGS EXCEPT WHERE SUCH ORDERS WOULD CONFLICT WITH THE FIRST LAW.

3 A ROBOT MUST PROTECT ITS OWN EXISTENCE AS LONG AS SUCH PROTECTION DOES NOT CONFLICT WITH THE FIRST OR SECOND LAWS.

Can Machines Think? Can They Feel?

Shall we treat the robots in our midst as our masters, our slaves or our partners? It's more a question for Hegel or Nietzsche than the technologists, but it's worth considering. The sociopathic superintelligences of science-fiction doomsday narratives are easy enough to dismiss; your laptop is not going to develop autocratic tendencies anytime soon. We could program robots to be our complacent slaves, but history shows that slavery dehumanizes the master along with the slave, hardly a happy prospect. As artificial intelligence becomes more humanoid, a kind of partnership seems the most likely outcome. Technology always seems to outrace policy, so why not consider diversity programs for non-human colleagues now?

Not so fast, say some AI experts. "Of course, people will experience robots as in some ways human, but I feel that applying concepts from human diversity will do a disservice to both," said Terry Winograd, a professor of computer science at Stanford University, and co-director of the Stanford Human-Computer Interaction Group. "The point of diversity training is learning to treat other people with the same respect as those of your type. I just don't believe that's applicable to computers."

Winograd said he does see the need for rules and procedures for working with artificial intelligences whose outputs have a direct impact on human life, such as a program that interprets financial and personal information to determine who qualifies for a home loan. "How you relate to machines that are making decisions in spheres that have a human outcome is a huge policy problem that needs to be worked on," he said. "The problem is the AI can't tell you what the criteria are. They don't have any introspection at all. They run the numbers, get an answer."

Winograd said that an AI's response to rudeness will simply be the one it is programmed to have, and it has no "feelings" to hurt. Robots are valuable assets, so they will require protection from abuse, but there will be a wide range of human-robotic interactions, and which ones will be considered appropriate will depend on circumstance. He sees a parallel with animal rights, in which some activists see any human use of animals as abusive, while other interest groups, such as laboratory scientists, apply something like a cost/benefit analysis to how non-human animals are treated. But neither rises to the level of a human-human interaction, he said.

Yoav Shoham, another Stanford computer scientist, teaches a freshman class called "Can Computers Think? Can They Feel?" At the beginning and end of the course, he polls students on those questions, noting the evolution in their answers. Along the way, they also explore questions regarding computers and free will, creativity, even consciousness, and their responses inevitably shift from predominantly noes to more yesses. Shoham readily concedes that he doesn't know the right answers, but adds that at least he knows that he doesn't know. He says his Socratic goal is to make the students doubt their automatic responses, and to at least start to question some of their biases. But he doesn't see a need to regulate human-robotic relations.

"I don't think that having Asimov rules for ethical treatment of robots is something that's needed now, any more than we need rules for our GPS or our smart watch," Shoham said. "It's well documented that we tend to anthropomorphize objects. I think there's a reason to be polite in communication with software, but not for that reason. I will occasionally cuss at my computer, and my wife is very upset by it. She doesn't like me to use that language, because I'll get used to it, and it will reflect on me. What if my kids hear me? We have social norms to use language in a certain way, and breaking it in one context will have spillover effects in different ways."

The Robot Told Me You Were Coming

William Gibson, the great science-fiction novelist, once told *The Economist*, "The future is already here—it's just not evenly distributed." Gibson has a knack for apt aphorisms, and that one hits the mark in Redmond, Wash., just outside Seattle. If you go there to visit Eric Horvitz, director of Microsoft Research, you will be greeted by an affable robot, which gives directions and uses casual language, like "No problem." Outside Horvitz's office you will meet Monica, his virtual admin, an attractive redheaded avatar with a British accent. "I was expecting you," she says. "The robot told me

you were coming." She might say that Horvitz is not in now, but she can schedule a meeting.

While you won't mistake the greeter or Monica for a human, they are personal and personable to a degree not commonly seen in robotics. After all, most robots in the modern enterprise are faceless mechanical muscle, performing one rote task day and night. They may be ruthlessly efficient, but there is not even a suggestion of personality or consciousness. The greeter and Monica are intentionally human-ish, providing a vision of a near future populated in part by well-informed, if not actually smart, artificial co-workers.

Horvitz has captured many hours of video showing that people usually speak to Monica in a polite way, even apologizing for misunderstood words and saying "thank you, nice to meet you," when they leave her. He says he thanks her, too, without thinking about it.

"We show the system working and the courtesy people have to it, then I walk up and the system recognizes me," Horvitz said. "The system always smiles when it sees me, and I tell folks that I enjoy the fact that it smiles only at me. These natural courtesies you extend to a system are pretty interesting. At the highest level, if we have systems that are anthropomorphic, like the kind, attractive British assistant at my door, and this system is working with multiple people, there are subtleties and nuances that make it an appropriate social actor in multiparty situations. The courtesy with which you treat the agent is the same you apply to other people."

While Monica cannot "think," and she really doesn't "know" anything, the system has been programmed with detailed data about Horvitz's schedule and priorities and has adequate analytical capability to make informed decisions about whom

to grant time with him, how soon and how much. Add to that a set of responsive facial expressions, natural language and instant interactive response, and it is easy to believe there is an intelligence at work. Horvitz and his team are creating a code base—software—that enables many forms of complex, layered interaction between machines and humans.

The system has situational awareness, meaning that it can take into account the physical space, people's comings and goings, their gestures and facial expressions, and the give and take of conversation between individuals. Horvitz said the goal is to build systems that can coordinate and collaborate with people in a fluid, natural manner. "I believe that systems will get so good, when you call up to work with a human entity on the phone, instead of someone saying to you 'this call may be recorded for quality assurance purposes,' it will say, 'I have to by law tell you that I am not a person.' When will that happen, when will that be important, is an interesting question to ask."

Microsoft Research is not a place for dreamers. The intent is to create technology for future products, and elements of Horvitz's work can already be encountered in Cortana, Microsoft's new virtual assistant for smartphones, which is named for the curvaceous AI heroine of the video game series Halo. Cortana is not as "nice" or knowledgeable as Monica, but she can set reminders, recognize a natural voice without the user having to input a predefined series of commands, and answer questions using information from Bing, Microsoft's search engine.

That sounds a lot like Siri, and it is, but at least Microsoft's engineers have given Cortana a sense of humor. The question, "Who's your daddy?" gets this response: "Technically speaking, that'd be Bill Gates. No big deal."

No Known Commercial Application, Yet...

While a visit to Redmond provides a vision of the near future, a journey to Reykjavík, Iceland, offers a glimpse of what's coming in a bit more distant tomorrow. A robotic agent, built by an international team led by researchers at Reykjavik University, is pushing the boundaries of artificial intelligence by automatically learning socio-communicative skills. The recently completed project, dubbed HUMANOBS, is not programmed in the conventional sense, but learns by observing and imitating humans in social situations.

"We essentially ditched all of engineering and computer science methodology wholesale, and approached it more from psychology and biology," said Kristin Thórisson, founding director of the Icelandic Institute for Intelligent Machines. "The starting point for those domains is really nature rather than mathematics. We set out a bunch of goals for the project that we thought we would achieve some of; we achieved all of them. The goal was to come up with an independent general learner that could be programmed by very high level goals."

In an ominous development for a still-working journalist, the HUMANOBS system's first achievement was to learn how to conduct an interview, simply by watching 20 hours of two humans in a mock-TV interview. Thórisson said that after just two or three minutes of observation, the system starts to understand what's going on, how to structure and conduct such an interview, and has generalized some of the main principles of human communication to a point that it can be asked to take over the interview, either in the role of the interviewer

or interviewee, and will continue to interact with the other person.

Thórisson said current workplace policies would suffice for interaction with learning systems, at least in the near term. "We have the etiquette for how we talk to our co-workers, and there's a lot of legal precedent," he said. "This technology is going to stir up that pot a little bit, but I think it's going to be very similar. If you don't want your co-workers to know something, don't tell your digital assistant. But if you think a bit further into the future, when the machines become harder to distinguish from humans, when you're at the point where you would feel a deep sense of loss if it got erased, like when you lose a pet or a loved one, then you might see something very different. But by then we'll see so many different things that I don't think this will be our main concern."

Pei Wang, the Temple University researcher, is developing systems similar to HUMANOBS, and he believes such machines will learn to differentiate between more and less pleasant human interactions. An AGI, or artificial general intelligence, he said, differs from an AI in that it initially knows nothing, but like a human baby, it is constantly learning. Like the baby, it is naïve, but it rapidly begins to evaluate the reliability of the many sources bombarding it with information. Those sources it deems more reliable will have greater influence, and in time, get better responses.

"In AGI, more and more people believe that emotion is a necessary aspect of high-level intelligence," Wang said. "It's nothing fancy. It will have a different attitude to other people or systems because of their relationship to it. If someone it likes makes a request, they will get more attention than others. I don't believe future AI will have emotions exactly like us, love and hate. But the basic motivation behind the emotion will be very simple: if you are polite, it will be polite. If you are mean, it will be mean."

When might we expect such systems? "If you force me to make a number, I will say something like 10 years," said Wang. "In the beginning, of course, it will be very simple. Five years is not enough, but I don't think we need 20. Because the basic principle is not magic. It's already understandable."

Machines of Loving Grace

Prognostications of emotional robots make some people profoundly uncomfortable, and it's not hard to see why. The science-fiction treatment of robots, which is the only one most of us know, has reliably veered between super-intelligent machines that will enslave humanity to seductive droids that will offer a pleasant but empty alternative to human intimacy. In her book "Alone Together," MIT psychologist Sherry Turkle frets about "sociable robots, which promise relationships where we will be in control, even if that means not being in control at all."

Will we have to worry about sexual harassment of AIs in the workplace? Perhaps. The author was astonished to be hit on while playing World of Warcraft using a female avatar. When the would-be suitor's advances were rebuffed, he—presumably it was a he—became abusive. Keep in mind that World of Warcraft avatars are two-dimensional and can only speak in text. An attractive artificial administrative assistant is bound to receive a certain amount of amorous attention. Is that good or bad? Who knows?

But best to be ready for it, sooner than later. "As these systems become conversational, our relationships with them will inevitably trend toward intimacy," said John Markoff, author of "Machines of Loving Grace: The Quest for Common Ground Between Humans and Robots," which will be published in August. People will rapidly accept closer relationships with robots, he said. "Today we are right where society was when ATMs were first introduced, when people refused to use them because they preferred to interact with a human teller. Almost overnight that preference reversed."

So if you're one of those people tempted to say, "you're welcome," when the ATM says, "thank you," don't worry. Your manners are just a little ahead of your time. ▰