

# Social Robots for the 21st Century

Daniel Ullrich and Andreas Butz, LMU Munich

*Two colleagues engage in a fictitious dialogue about the future of robotics.*

**H**ello Andreas, I heard you have some outdated opinions about the future of robotics and thought I would help you out!

Advances in human–robot interaction will clearly bring many benefits to society. Consider residential care facilities, where more and more people spend their retirement. Typically in such facilities, staff are occupied with basic nursing chores and social aspects are almost totally neglected. Robots could assume several repetitive manual duties—cleaning, fetching items, bathing patients. Indeed, companies are already developing such robots. This would enable human caregivers to spend more time chatting with patients and addressing their psychological needs.

Advanced robots could even take over many social tasks in nursing homes. Therapeutic robots such as PARO ([www.parorobots.com](http://www.parorobots.com)) and JustoCat ([www.justocat.com](http://www.justocat.com)) already demonstrate that they can improve patients' general relaxation as well as facilitate their socialization with one another and with caregivers.

Future social robots will be able to substantially support humans in many areas beyond caregiving: companionship,

entertainment, security, transportation, education, customer service, personal assistance, sales, tourist guidance—the list goes on.

What's your take on the future of robotics?  
Daniel

## ROBOT IDENTITY CRISIS

Dear Daniel, Glad you asked! Let me tell you why social robots are just another overhyped technology, and will fail to deliver on their grandiose promises.

Let's establish some common ground first. Robots are machines, right? I hope you're not one of those people who refer to a robot as "he" or "she." I can't imagine developing human sentiments toward a robot, since I know that everything I tell it won't evoke true feelings but will simply be analyzed algorithmically. When was the last time you had a trusting relationship with a machine? That worked out well in *2001* and *Ex Machina*, didn't it :-)?

Besides, deep AI will never happen! I've been in this field too long. The only instances of true AI are highly specialized or occur in a sandbox; deep learning is mostly



wishful thinking by lay journalists or researchers seeking publicity. I just hope that the social robot bubble will burst before all “obsolete” humans have been fired.

Live long and prosper,  
Andreas

Dear Andreas,

I agree that robots are more or less autonomous software-controlled machines, and their human-like shape implies they can do things they actually can't.

I wouldn't refer to my robot as “he,” nor would I talk to him directly. But people can still form relationships with, and have feelings about, technology. For example, a participant in a recent experimental evaluation of SleepCare, an automated e-coach designed to help users overcome insomnia, reported of forming “a ‘bond’ with my coach, although I am very much aware that it's just an algorithm.”<sup>1</sup> Requiring robots to possess human qualities like consciousness, empathy, and sensibility is demanding more than is necessary.

Wouldn't you agree that building a perfect human-like robot is overshooting the mark?

Peace and long life,  
Daniel

Dear Daniel,

Ha ha, I got you! You wrote that you wouldn't refer to your robot as “he,” but in the same sentence used “him” instead of “it.” Is that a Freudian slip?

I agree that people can have personal relationships with their pets. People likewise might develop feelings for robots, but they'd only be projecting their own concepts of intelligence and personality into the machine.

By “demanding more than is necessary,” do you imply that confirming people's projections about human-like robots is sufficient? Need we only

push AI far enough to make it indistinguishable from magic—apparent rather than true intelligence? If so, that doesn't shed a very flattering light on the human mind and how easily it can be deceived.

Can you give me a convincing example?

Cheers,

A slightly unsettled Andreas

### MISERABLE MURPHY

Dear Andreas,

You caught me ;-), but I'm not alone: all of my students refer to robots using a

(see Figure 1). Before seeing the doctor, many children are stressed or scared and lack the coping strategies that adults have acquired. Parents often are of little help because they themselves aren't patients and “miserery loves only miserable company.”<sup>3</sup> As a pint-size “patient” with health concerns also waiting for the doctor, Murphy provides such “miserable company.” Equipped with face and speech recognition, he [sic!] interacts with the children by sharing feelings and stories about the silly accidents that happened to him and by moving the children's cognitive focus away

Requesting robots to possess human qualities like consciousness, empathy, and sensibility is demanding more than is necessary.

personal pronoun. I didn't realize this before and wonder when they switched from “it” to “he.” Actually, treating a robot like a human is understandable if enough parameters are human-like—for example, intelligence, sensitivity, anthropoid appearance.

I fully agree with your explanation for this: projection. Humans are masters of pattern recognition. Our brains—neural networks—work top-down and bottom-up simultaneously: we recognize, for example, shapes by visual sensation and classify them by our knowledge of the world. Therefore, providing a robot with appropriate human characteristics will trigger our brains to classify it as an animate intelligent being and provoke associated emotional reactions.

You asked for an example. My research group has developed a concept for a social robot called Miserable Murphy<sup>2</sup> to comfort anxious children in a pediatric waiting room

from worrying about the impending doctor visit and by asserting their own strengths and capabilities in coping with illness.

We haven't tested Murphy with real patients, but reactions to our video prototype are promising.<sup>2</sup> Child subjects appreciated the robot's company in an anxious situation and rated his fairly limited verbal dialogue capabilities as stimulating. Adults who watched a video of the interactions (<https://youtu.be/seb9ZefGZmk>) were amazed by Murphy's seemingly intelligent reactions, oblivious to the fact that the children weren't real patients and the robot's reactions were scripted.

The positive responses to Murphy reinforce your point about projection: because the limits of science are unknown, people don't know what to expect from technology. If they're ready to believe in intelligent robots, we might as well use this to our advantage.



Figure 1. Miserable Murphy provides comfort to anxious children waiting to see the doctor.

Do you think this is ethically questionable?

Daniel, who's curious where all this is going

**ETHICAL DILEMMAS**

Hi Daniel,

I admit that in this and similar situations, robots like Murphy could actually help real children on an emotional level. The psychological foundations behind this idea make it really promising, and I could see my kids going for it.

Nevertheless, I foresee a societal problem. You said that we project human intelligence and sensitivity into a robot who (I mean *that!*) sends us the right signals. If someday our world was filled with robots, we could find ourselves surrounded by humanoid entities that can be repaired if they break and discarded if we get bored or want an upgrade, and can't be humiliated or appalled by our bad behavior. Over time, across generations, we might start to treat people like machines—carelessly or, worse, abusively. By making robots lifelike and

encouraging people to project human qualities into robots, don't we risk blurring the line between people and machines too much?

Cheers,  
Andreas

Dear Andreas,

I assume that human-robot interaction (HRI) will affect social interaction in general, but differently: HRI advances will lead to us treating robots like humans, not vice versa. I doubt that robots will become ubiquitous before reaching a certain stage of development. The breakthrough will occur when they either really are intelligent enough to be indistinguishable from humans or convincingly use our human heuristics and stereotypes to at least appear intelligent.

I see two ethical dilemmas in future robotics applications.

First, how should robots behave ethically? What should be the rationale for their decisions? Isaac Asimov's Three Laws of Robotics<sup>4</sup> are insufficient. In a recent study replicating a variation of the trolley problem,<sup>5</sup> human participants rated the choices of human

versus robot agents faced with doing nothing and letting a runaway trolley kill five people ahead on the main track or intervening to redirect the trolley to a side track and killing one person. Robot agents that didn't choose the utilitarian option—that is, sacrificing one person to save five—were blamed for making the morally wrong decision much more frequently than human agents. Wouldn't you agree that robots are expected to behave this way—analytically, logically, machine-like?

Second, how should humans treat robots? What's the minimum requirement for being "more" than a machine? Is intelligence the only criterion, or must a robot demonstrate some other essential human quality?

Your overly optimistic colleague,  
Daniel

Dear Daniel,

I'm not a philosopher: I have opinions about robot ethics but might not be able to back them up scientifically, so please bear with my makeshift approach.

What would it take for me to see robots as more than machines? Frankly,

flesh and blood. I can't imagine ever equating robots with humans, no matter how indistinguishable they become. Of course, I might be dazzled by sophisticated robots' human-like charms or intelligence, though that's not necessarily a good thing—recall the doomed protagonist of *Ex Machina* who was emotionally manipulated into freeing a homicidal robot. You're the specialist on humans, so I'll have to take your word for it that someday we'll come to accept robots as more than just machines. But when looking at AI research, I don't see this happening anytime soon.

With respect to your first question, I basically agree with utilitarianism, but I don't think it's always right. In *Star Trek II*, Spock asserted that “the needs of many outweigh the needs of the few,” but in *Star Trek III*, Kirk, after saving Spock, points out that “sometimes the needs of the one outweigh the needs of the many.” You can guess with whom my sympathies lie. Perhaps this should be the moral basis of a mixed human-robot society: let all robots be utilitarian, leaving humans to show feelings and put friendship above utility and rhyme over reason.

I guess that answers your other question. Not only do I believe that robots will never develop true feelings such as friendship, I don't want them to. Perhaps I'm old-fashioned, a technological pessimist who subscribes to Douglas Hofstadter's argument that because no formal system will ever fully understand itself, human intelligence will never exist in manmade machines.<sup>6</sup> I don't care how indistinguishable humans and robots become. There is no magic!

Your stubborn colleague,  
Andreas

P.S. A considerable part of society believes in magic, so perhaps merely the appearance of magic is enough. Some people firmly accept religion, some reject it, and others claim that religion is a useful and understandable encoding of morality. Should we

build “religious robots” so that people can form a consistent mental model of their ethics?

## DO YOU BELIEVE IN MAGIC?

Dear Andreas,

Finally, we're in the realm of magic ;-)! Arthur C. Clarke's third law—“any sufficiently advanced technology is indistinguishable from magic”<sup>7</sup>—makes sense in psychological terms: people strive for explanations to make sense of their reality. For example, without knowledge of magnetism you would struggle to explain objects levitated by this physical principle, and “magic” would be a simple explanation.

Your closeness to AI research might account for your skepticism about the

While we can fully describe a single nerve cell, we can't fully understand how the entire nervous system works. Yet we can use the same principle to build working neural networks, and there's no logical reason to prevent us from combining multiple networks and increasing complexity to a level comparable to that of human brains.

Now, as to your other question: how should advanced robots behave, assuming that they can be made in the first place? If robots were limited to utilitarian behavior, how could they possibly interact with humans without such encounters leading to one misunderstanding after another? Robots must have a basic understanding of feelings, humor, emotions, irrationality, and all the other qual-

---

**Robots must have a basic understanding of feelings, humor, emotions, irrationality, and all the other qualities that make us human.**

future role of social robots: your intimate knowledge about the technology's current limitations leaves no room for magic and imagination, two factors that make interaction with robots much easier, at least for the robots ;-). We both must acknowledge that there are widely differing opinions on what constitutes desirable HRI. Several companies are developing human-like robotic dolls made of silicone to function as “romantic” partners for their owner. For me, these are located at the very bottom of the “uncanny valley,”<sup>8</sup> but they have their paying customers.

You brought up Hofstadter's argument that no formal system will ever fully understand itself. While I agree with this, I disagree with your conclusion that “human intelligence will never exist in man-made machines.” This assumes that all inventions need to be fully understood, which I don't believe. Consider neural networks.

ities that make us human. Why not give their personalities these human “qualities”? I would love to have a sarcastic robot hanging around in my apartment.

After all, there probably won't be a single type of robot. The question is whether we should prevent robots from entering particular domains that should be reserved for humans. Ironically, many emerging robotics applications are in a very social domain: caregiving. Thus, this question might already be answered 8-(-.

You finished with thoughts about belief and religion. I'm not religious, nor do I believe in magic—which is ironic given my arguments for social robots with a “magical” appearance ;-). But I can see that religious faith or belief in magic can increase well-being and happiness. Let robots utilize this link, as long as it does no harm.

So long,  
Daniel

**D**ear Daniel, Agreed, with one exception: you assert that by combining enough sufficiently complex neural networks we might eventually obtain intelligence. For me, that's like saying we could eventually obtain life by combining enough oxygen, hydrogen, and carbon at the right temperatures. Given entropy, I doubt this, though I can't disprove it.

Maybe we should leave it at that and agree that our speculations about social robots are just that—speculations. Fun, but ultimately futile. We'll have to wait and see. Perhaps we'll be fortunate enough to shape the future of robotics the way we want it to be. Let's meet next week and brainstorm about some grant proposals ;-)!

Cheers,  
Andreas

### REFERENCES

1. R.J. Beun et al., "Improving Adherence in Automated E-Coaching," *Persuasive Technology*, LNCS 9638, Springer, 2016, pp. 276–287.
2. D. Ullrich, S. Diefenbach, and A. Butz, "Murphy Miserable Robot," *Proc. CHI Conf. Extended Abstracts on Human Factors in Computing Systems (CHI EA 16)*, 2016, pp. 3234–3240.
3. S. Schachter, *The Psychology of Affiliation*, Academic Press, 1959.
4. I. Asimov, *I, Robot*, Gnome Press, 1950.
5. B.F. Malle et al., "Sacrifice One for the Good of Many? People Apply Different Moral Norms to Human and Robot Agents," *Proc. 10th Ann. ACM/IEEE Int'l Conf. Human-Robot Interaction (HRI 15)*, 2015, pp. 117–124.
6. D.R., Hofstadter, *Gödel, Escher, Bach*, Basic Books, 1979.
7. A.C. Clarke, "Hazards of Prophecy: The Failure of Imagination," *Profiles of the Future*, Harper and Row, 1962, pp. 12–21.
8. M. Mori, "The Uncanny Valley" (orig. pub. 1970), K.F. MacDorman and N. Kagei, trans., *IEEE Robotics & Automation Mag.*, vol. 19, no. 2, 2012, pp. 98–100.

**DANIEL ULLRICH** is a research assistant in the Human-Computer Interaction group in the Institute for Informatics at LMU (Ludwig Maximilians University) Munich. Contact him at [daniel.ullrich@ifi.lmu.de](mailto:daniel.ullrich@ifi.lmu.de).

**ANDREAS BUTZ** leads the Human-Computer Interaction group in the Institute for Informatics at LMU Munich. Contact him at [andreas.butz@ifi.lmu.de](mailto:andreas.butz@ifi.lmu.de).

Keeping  
**YOU** at the  
Center  
of Technology  
**myCS**



Publications your way,  
when you want them.

The future of publication delivery  
is now. Check out myCS today!

- **Mobile-friendly**—Looks great on any device—mobile, tablet, laptop, or desktop
- **Customizable**—Whatever your e-reader lets you do, you can do on myCS
- **Personal Archive**—Save all your issues and search or retrieve them quickly on your personal myCS site.

Stay relevant with the IEEE Computer Society

More at [www.computer.org/myCS](http://www.computer.org/myCS)

IEEE  computer society