

Introduction: Open Questions in Roboethics

John P. Sullins

Published online: 12 August 2011
© Springer-Verlag 2011

1 The Rise of the Robotethicist

Roboethics is the recent offshoot of computer ethics that pays special attention to the alterations that need to be made to computer ethics when we give the computer mobility and a means to interact directly in the human environment. The closely related field of machine morality explores how ethical systems and behaviors may be programmed into social robotics applications. As robots move from the factory floor into our homes and work lives, they stand to change key aspects of the way our lives are lived. In order to be successful, these machines must be programmed with the ability to navigate the human life world without committing ethical faux pas or moral outrage. Thus, the roboethicist is tasked not only with critiquing the attempts of robots engineers to achieve the integration of these machines into our life world, but also, and more importantly, with suggesting means of achieving better results than what is presently on offer.

The undeniable roots of roboethics begin in the world of science fiction. The very coining of the word “robot” in Karel Čapek’s 1936 play, *RUR*, is loaded with ethical import. The Czech word “Robota” refers to labor or servitude, which gives us the uncomfortable inference that roboethics refers to a kind of slave ethics. I reject this connotation and it is just an unfortunate byproduct from the literary trope of the robot rebellion that Čapek began with his play and Fritz Lang masterfully solidified in the human psyche with his film *Metropolis*, something which Hollywood has been reiterating ever since. There is no need to reenact this unfortunate future in reality. As the great science fiction writer Philip K. Dick once observed, the duty of science fiction is to imagine dystopian futures so that we don’t actually have to live them. With this in mind, we can then see that the job of the roboethicist is not simply science fiction, it is instead to help avoid the imagined robo-apocalypse and help build an alternative future where robots are not resentful slaves or out of control killing machines, but instead more like pets and perhaps someday even friends or possibly, in the very far future, even colleagues. In the near future, the job of the

J. P. Sullins (✉)

Sonoma State University, 1801 East Cotati Aveune, Rohnert Park, CA 94928, USA
e-mail: John.sullins@sonoma.edu

roboethicist is to ensure that we do not harm each other too deeply with these machines as they grow in complexity and capability.

Gianmarco Veruggio seems to have coined the term “roboethics” in 2002 at the first roboethics workshop organized around an IEEE robotics conference.¹ At that time, it was decided to separate the field into two allied subfields. One is machine ethics or machine morality, which is concerned with describing how machines could behave ethically towards humans. The other is roboethics itself, which is concerned with how humans relate to these machines in both the design and use phase of their operation. In the last 9 years though, these terms have drifted a bit and you will hear expressions such as “machine ethics,” “machine morality,” “roboethics,” “robot ethics,” and “moral machines” all used somewhat synonymously to refer to the ethical concerns raised by robotics technologies.

Since that time there have been numerous articles printed, workshops and conference tracks organized, special issues of journals, blogs and Facebook groups formed, as well as a few important book projects. But there is much left to do and it is my purpose here to try to interest more people to join this growing area of research. Robotics technology’s move into the home is roughly where the personal computer was in the 1970’s. If such trend continues, then we can expect personal robotics and military robotics to move quickly into the home, workplace and battlefield. It is therefore our duty to stay ahead of that curve in order to anticipate and help alleviate the ethical impacts of these technologies.

One further conceptual complexity needs to be mentioned here as well. Robots come in two broad categories, autonomous and non-autonomous. Roughly speaking, “autonomy” typically refers to the level of human control and oversight over the robot’s actions and decisions. When one speaks of “autonomous robots,” one is generally not making any strong claim regarding the philosophical free will of the machine. It is simply the acknowledgement that autonomous robots make the majority of their decisions using computational systems, whereas non-autonomous or telerobots have at least some human oversight and input into the decisions they make.²

While the media has us all used to the idea of autonomous robots, as it turns out they are very difficult to make and so the robots we see in use today are all largely telerobots. Therefore roboethicists should focus a bit more on how telerobots alter the ethical thinking of their users since machines making autonomous ethical decisions are still only a theoretical possibility.

2 Open Questions in Roboethics

As roboethics is a young field of study, there are many interesting open questions and subfields of study. My list here is not meant to be exhaustive but it is what I believe to be the most interesting at this time.

¹ It is possible that the term was used prior to this date but the 2002 workshop is certainly where the term became more widely used.

² There is more complexity in the case of machines that house the majority of their artificial intelligence capabilities in a remote machine and then teleoperate a drone or drones of some type. I will gloss over this here and just treat these machine architectures as autonomous robots for now.

2.1 Military Applications

This is by far the most important of the subfields of roboethics. It would have been preferable had we worked through all the problems of programming a robot to think and act ethically before we had them make life and death decisions, but it looks like that is not to be. While teleoperated weapons systems have been used experimentally since the Second World War, there are now thousands of robotic weapons systems deployed all over the world in every advanced military organization and in an ad hoc way by rebel forces in the Middle East (Singer 2009). Some of the primary ethical issues to be address here revolve around the application of just war theory. Can these weapons be used ethically by programing rules of warfare, the law of war and just war theory into the machine itself? Perhaps machines so programmed would make the battlefield a much more ethically constrained space? How should they be built and programmed to help war fighters make sound and ethical decisions on the battlefield? Do they lower the bar to entry into conflict too low? Will politicians see them as easy ways to wage covert wars on a nearly continuous level? In an effort to keep the soldier away from harm, will we in fact bring the war to our own front door as soldiers telecommute to the battlefield? What happens as these systems become more autonomous? Is it reasonable to claim that humans will always be “in” or “on the loop” as a robot decides to use lethal force?

2.2 Privacy

Robots need data to operate. In the course of collecting, data they will collect some that people may not want shared but which the machine needs nonetheless to operate. There will be many tricky conundrums that have to be solved as more and more home robotics applications evolve. For instance, if we imagine a general-purpose household robot of the reasonably near future, how much data of the family’s day-to-day life should it store? Who owns that data? Might that data be used in divorce or custody settlements? Will the robot be another entry for directed marketing to enter the home?

2.3 Robotic Ethical Awareness

How does a machine determine if it is in an ethically charged situation? And assuming it can deal with that problem, which ethical system should it use to help make its decision? Philosophers such as John Dewey and later Mario Bunge have argued that a technology of ethics is possible and in some ways preferable (Sullins 2009). I am certain they were not thinking of robots when they made these arguments, but their view that ethics is transactional and instrumental allow us to extend their ideas to the claim that ethics is computational. Thus, it is not out of the question that machine ethics is possible. Yet we are sorely lacking on the specifics needed to make any of these claims anything more than theoretical. Engineers are wonderfully opportunistic and do not tend to have emotional commitments to this or that school of thought in ethics. Therefore, what we see occurring today is that they tend to make a pastiche of the ethical theories that are on offer in philosophy and pick and choose the aspects of each theory that seem to work and deliver real results.

2.4 Affective Robotics

Personal robots need to be able to act in a friendly and inviting way. This field is often called social robotics, sociable robotics, or affective computing, and was largely the brainchild of Cynthia Breazeal, from the Massachusetts Institute of Technology (MIT) robotics lab (2002). The interesting ethical question here is: if your robot acts like your friend, is it really your friend? Perhaps that distinction does not even matter? With sociable robotics, the machine looks for subtle clues gathered from facial expression, body language, perhaps heat signatures or other biometrics and uses this data to ascertain the user's emotional state. The machine then alters its behavior to suit the emotional situation and hopefully make the user feel more comfortable with the machine. If we come to accept this simulacrum of friendship, will this degrade our ability to form friendship with other humans? We might begin to prefer the company of machines.

2.5 Sex Tobots

It seems strange but it is true that there are already semiresponsive sex dolls that do count as a minor type of robot. These machines are such a tantalizing dream for some roboticists that there is little doubt that this industry will continue to grow. This category of robotics supercharges the worries raised by affective robotics and adds a few more. Sociable robots examine the user biometrics so the robot can elicit friendly relations, but here the robot examines biometrics to elicit sexual relations. A sex robot is manipulating very strong emotions and if we thought video games were addictive, then imagine what kind of behavior might be produced by a game consul with which one could have sex. These machines are likely to remain on the fringe of society for some time, but the roboticist David Levy has argued that since this technology can fulfill so many of our dreams and desires, it is inevitable that it will make deep market penetration and eventually will be widespread in our society (Levy 2007). This will result in many situations that will run the spectrum from tragic, to sad, to humorous. The key point here is: whether the machines can really be filled with love and grace or whether we are just fooling ourselves with incredibly expensive and expressive love dolls. I can easily grant that engineers can build a machine with which many would like to have sex, but can they build a machine that delivers the erotic in a philosophical sense? Can they build a machine that can make us a better person for having made love to it?

2.6 Carebots

Somewhat related to the above are carebots. These machines are meant to provide primary or secondary care to children, the elderly and medical patients. There are already a number of these machines, such as the Paro robot, in service around the world. On one end of the scale, one has something like Paro, a robot that is meant to provide artificial pet therapy for its users. Towards the middle of the scale, one would have machines built to assist medical caregivers in lifting and moving patients or helping to monitor their medications or just to check in with patients during their stay. At the far end of the scale, one would have autonomous or semi-autonomous machines that would have nearly full responsibility in looking after children or the elderly in a home setting. Here again, we have some of the same issues raised by social robotics and the

concomitant privacy issues. But in addition to those you have the troubling problem of why other humans are not taking care of their own children and elderly. What kind of society are we creating where we wish to outsource these important human relations to a machine allowing younger generations to simply ignore the elderly?

2.7 Medibots

These are related to carebots but I am specifically thinking here of robots that assist in surgery and other life and death medical practices such as administering medication. Often, the surgeons using these machines are close by the operating theater, but this technology is also used to allow a surgeon to work on a patient many thousands of miles away. This technology can be useful when dealing with a wounded soldier on a distant battlefield or a patient with serious conditions who is living in remote or economically depressed places of the world. This technology puts a new wrinkle on many of the standard medical ethics issues and we need more medical ethicists to study this phenomenon in depth.

2.8 Autonomous Vehicles

Our roadways could change in a very radical way. Autos and large transportation vehicles of the near future may have no human driver. Already some luxury vehicles will take over in emergency breaking situations or when the driver falls asleep at the wheel. A number of autos will park themselves completely autonomously. The vast majority of the ethical issues involved here will be legal in nature, but there will also be issues of trust involved. For instance, can one trust a vehicle to make the right decisions when those decisions mean the lives of you, your family and all those around you? There have already been deaths caused by faulty automatic navigation services because people robotically follow the directions of the GPS machine no matter what it says, even if it is giving incorrect directions that lead one into dangerous situations.

2.9 Attribution of Moral Blame

This is one of the biggest conundrums in roboethics. Nearly all moral systems have some way of assessing which moral agent involved in a system is to blame when things go wrong. Most humans respond to blame and punishment and might modify their behavior to avoid it when possible. But how does one blame a machine? Will people use robots as proxies for the bad behavior in order to remove themselves from blame? When a military robot kills an innocent civilian, who is to blame? If you are asleep in your robotic car and it runs down a pedestrian, did you commit manslaughter or are you just an innocent bystander?

2.10 Environmental Robotics

There are two ways to look at the environmental ethics impacts of robotics. One is to look at the impact of the manufacture, use and disposal of robots. Currently, there is no green robotics movement and we should push for this to be developed. A second

interesting idea is that robotics could provide an invaluable tool for gathering data about environmental change. The very same robots that are used to monitor enemy troops and scour the ocean floor for enemy activity can be easily re-tasked to monitor forests, ocean ecosystems, protect whale and dolphins or any number of environmental tasks that unaided humans find difficult.

3 Robotics, War and Peace

This special issue is an attempt to advance our understanding of the many issues raised above. The articles collected here represent some of the very best thought on these subjects. I would like to thank the many referees who worked on this project; their unsung efforts ensured that this issue is a valuable contribution to the growing scholarship on the ethical impacts of robotics technology. Robotics and warfare dominates the conversation at this time but it is my sincere hope that the many conflicts that plague our world can diminish and we can move to the study of more peaceful applications of this fascinating technology.

References

- Breazeal, C. (2002). *Designing sociable robots*. Cambridge, MA: The MIT Press.
- Levy, D. (2007). *Love and sex with robots, the evolution of human-robot relationships*. Harper/HarperCollins: New York, USA.
- Singer, P. W. (2009). *Wired for war*. New York: Penguin Press HC.
- Sullins, J. (2009). Artificial moral agency in technoethics. In R. Luppincini & R. Adell (Eds.), *Handbook of research on technoethics*. New York: IGI Global.