Artificial Intelligence 175 (2011) 942–949

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Robot ethics: Mapping the issues for a mechanized world

Patrick Lin [a,*,1], Keith Abney [b,c,2], George Bekey [a,3]

[a] California Polytechnic State University, Philosophy Department, 1 Grand Avenue, San Luis Obispo, CA 93407, USA
[b] California Polytechnic State University, College of Engineering, 1 Grand Avenue, San Luis Obispo, CA 93407, USA
[c] University of Southern California, Viterbi School of Engineering, Los Angeles, CA 90089-0781, USA

ARTICLE INFO

Article history:
Received 13 September 2010
Accepted 28 November 2010
Available online 19 January 2011

Keywords:
Robot
Robotics
Ethics
Society
Philosophy
Psychology
Law
Policy
Safety
Error

ABSTRACT

As with other emerging technologies, advanced robotics brings with it new ethical and policy challenges. This paper will describe the flourishing role of robots in society—from security to sex—and survey the numerous ethical and social issues, which we locate in three broad categories: safety & errors, law & ethics, and social impact. We discuss many of these issues in greater detail in our forthcoming edited volume on robot ethics from MIT Press.

Bill Gates recently observed that “the emergence of the robotics industry … is developing in much the same way that the computer business did 30 years ago” [18]. As a key architect of the computer industry, his prediction has special weight. In a few decades—or sooner, given exponential progress forecasted by Moore’s Law—robots in society will be as ubiquitous as computers are today, he believes; and we would be hard-pressed to find an expert who disagrees.

But consider just a few of the challenges linked to computers in the last 30 years: They have displaced or severely threatened entire industries, for instance, typewriter manufacturing and sales by word-processing software, accountants by spreadsheets, artists by graphic-design programs, and many local businesses by Internet retailers. Customer-tracking websites, street-view maps, and the free and anonymous flow of information online still raise privacy concerns. The digital medium enables sharing that may infringe on copyright claims, and a largely unregulated process of registering domain names has led to charges of cybersquatting or trademark disputes. The effects of social networking and virtual reality on real-world relationships are still unclear, and cyberbullying is a new worry for parents. Internet addiction, especially to online gaming and pornography, continues to ruin real lives. Security efforts to protect corporate networks and personal computers require a massive educational campaign, not unlike safe-sex programs in the physical world. And so on.

* Corresponding author. Tel.: +1 (805) 756 2041.
E-mail addresses: palin@calpoly.edu (P. Lin), kabney@calpoly.edu (K. Abney), gbekey@calpoly.edu (G. Bekey).

[1] Patrick Lin is the director of the Ethics + Emerging Sciences Group—http://ethics.calpoly.edu—and an assistant philosophy professor at California Polytechnic State University, San Luis Obispo. He is also an affiliated scholar at Stanford Law School’s Center for Internet and Society and a visiting research fellow at Australia’s Centre for Applied Philosophy and Public Ethics (CAPPE).

[2] Keith Abney is a senior philosophy lecturer and bioethicist at California Polytechnic State University, San Luis Obispo. Tel.: +1 (805) 756 2041.

[3] George Bekey is a professor emeritus of computer science, electrical engineering, and biomedical engineering at University of Southern California as well as founder of its robotics lab. He is also a distinguished adjunct professor of engineering at California Polytechnic State University, San Luis Obispo. Tel.: +1 (805) 756 2131.

0004-3702/$ – see front matter
doi:10.1016/j.artint.2010.11.026

To be clear, these are not arguments that the computer industry should never have been developed, but only that its benefits need to be weighed against its negative effects. However, we are not interested in making such a cost-benefit evaluation here but would like to focus on an important lesson: If the evolution of the robotics industry is analogous to that of computers, then we can expect important social and ethical challenges to rise from robotics as well, and attending to them sooner rather than later will likely help mitigate those negative consequences.

Society has long been concerned with the impact of robotics, even before the technology was viable. Beginning with the first time the word 'robot' was coined [13], most literary works about robots are cautionary tales about insufficient programming, emergent behavior, errors, and other issues that make robots unpredictable and potentially dangerous (e.g., [5,6,16,45]). In popular culture, films continue to dramatize and demonize robots, such as *Metropolis*, *Star Wars*, *Blade Runner*, *Terminator*, *AI*, and *I, Robot*, to name just a few. Headlines today also stoke fears about robots wreaking havoc on the battlefield as well as financial trading markets, perhaps justifiably so (e.g., [19]).

A loose band of scholars worldwide has been researching issues in robot ethics for some time (e.g., [42]). And a few reports and books are trickling into the marketplace (e.g., [43,26,38]). But there has not yet been a single, accessible resource that draws together such thinking on a wide range of issues, e.g., programming design, military affairs, law, privacy, religion, healthcare, sex, psychology, robot rights, and more. To fill that need, the authors of this paper are in the process of editing a collection of robot-ethics papers [28] for MIT Press, a leading publisher in robotics as well as ethics. In this journal paper, we will briefly introduce the major issues in robot ethics.

1. What is a robot?

Let us start with a basic issue: What is a robot? Given society's long fascination with robotics, it seems hardly worth asking the question, as the answer surely must be obvious. On the contrary, there is still a lack of consensus among roboticists on how they define the object of their craft. For instance, an intuitive definition could be that a robot is merely a computer with sensors and actuators that allow it to interact with the external world; however, any computer that is connected to a printer or can eject a CD might qualify as a robot under that definition, yet few roboticists would defend that implication.

Certainly, artificial intelligence (AI) by itself can raise interesting issues, such as whether we ought to have humans in the loop more in critical systems, e.g., those controlling energy grids and making financial trades, lest we risk widespread blackouts and stock-market crashes [14,29]. But robots or embodied AI that can directly exert influence on the world seem to pose additional or special risks and ethical quandaries we want to distinguish here. A plausible definition, therefore, needs to be more precise and distinguish robots from mere computers and other devices.

We do not presume we can resolve this great debate here, but it is important that we offer a working definition prior to laying out the landscape of current and predicted applications of robotics. In its most basic sense, we define "robot" as *an engineered machine that senses, thinks, and acts*: "Thus a robot must have sensors, processing ability that emulates some aspects of cognition, and actuators. Sensors are needed to obtain information from the environment. Reactive behaviors (like the stretch reflex in humans) do not require any deep cognitive ability, but on-board intelligence is necessary if the robot is to perform significant tasks autonomously, and actuation is needed to enable the robot to exert forces upon the environment. Generally, these forces will result in motion of the entire robot or one of its elements (such as an arm, a leg, or a wheel)" [9].

This definition does not imply that a robot must be electromechanical; it leaves open the possibility of biological robots, as well as virtual or software ones. But it does rule out as robots any *fully* remote-controlled machines, since those devices do not "think", e.g., many animatronics and children's toys. That is, most of these toys do not make decisions for themselves; they depend on human input or an outside actor. Rather, the generally accepted idea of a robot depends critically on the notion that it exhibits some degree of autonomy or can "think" for itself, making its own decisions to act upon the environment. Thus, the US Air Force's Predator unmanned aerial vehicle (UAV), though mostly tele-operated by humans, makes some navigational decisions on its own and therefore would count as a robot. By the same definition, the following things are not robots: conventional landmines, toasters, adding machines, coffee makers, and other ordinary devices.

As should be clear by now, the definition of "robot" also trades on the notion of "think", another source of contention which we cannot fully engage here. By "think", what we mean is that the machine is able to process information from sensors and other sources, such as an internal set of rules either programmed or learned, and to make some decisions autonomously. Of course, this definition merely postpones our task and invites another question: What does it mean for machines to have autonomy? If we may simply stipulate it here, we define "autonomy" in robots as *the capacity to operate in the real-world environment without any form of external control, once the machine is activated and at least in some areas of operation, for extended periods of time* [9].

Thus again, *fully* remote- or tele-operated machines would not count as autonomous, since they depend on external control; they cannot "think" and therefore cannot act for themselves. But what about the everyday desktop or laptop computers: Are they autonomous? Doesn't their programming count as human inputs or external control in some important sense? If so, how can robots ever be said to be free from external control, if all robots are computers (electromechanical or otherwise) at their core?

These are all good questions that demand answers, for a complete discussion of what it means to be a robot. Many will engage other difficult issues from technical to philosophical, such as complexity, unpredictability, determinism, responsibility, and free will. As such, we cannot offer a complete discussion given space limitations of this paper, and we will have to content ourselves with the working definitions stipulated above—which should be enough to understand why we include some machines and not others in the following section.

2. Robots today and tomorrow

Robots are often tasked to perform the "three Ds", that is, jobs that are dull, dirty, or dangerous. For instance, automobile factory robots execute the same, repetitive assemblies over and over, with precision and without complaint; military surveillance UAVs patrol the skies for far more hours than a human pilot can endure at a time. Robots crawl around in dark sewers, inspecting pipes for leaks and cracks, as well as do the dirty work in our homes, such as vacuuming floors. Not afraid of danger, they also explore volcanoes and clean up contaminated sites, in addition to more popular service in defusing bombs and mediating hostage crises.

We can also think of robots more simply and broadly—as human replacements. More than mere tools which cannot think and act independently, robots are able to serve in many old and new roles in society that are often handicapped, or made impossible, by human frailties and limitations; that is, semi- and fully-autonomous machines could carry out those jobs more optimally. Beyond the usual "three D's", robots perform delicate and difficult surgeries, which are risky with shaky human hands. They can navigate inaccessible places, such as the ocean floor or Mars. As the embodiment of AI, they are more suited for jobs that demand information processing and action too quick for a human, such as the US Navy's Phalanx CIWS that detects, identifies, and shoots down enemy missiles rapidly closing in on a ship. Some argue that robots could replace humans in situations where emotions are liabilities, such as battlefield robots that do not feel anger, hatred, cowardice, or fear—human weaknesses that often cause wartime abuses and crimes by human soldiers [4]. Given such capabilities, we find robots already in society or under development in a wide range of roles, such as:

Labor and services: Nearly half of the world's 7-million-plus service robots are Roomba vacuum cleaners [21], but others exist that mow lawns, wash floors, iron clothes, move objects from room to room, and other chores around the home. Robots have been employed in manufacturing for decades, particularly in auto factories, but they are also used in warehouses, movie sets, electronics manufacturing, food production, printing, fabrication, and many other industries.

Military and security: Grabbing headlines are war robots with fierce names such as Predator, Reaper, Big Dog, Crusher, Harpy, BEAR, Global Hawk, Dragon Runner, and more. They perform a range of duties, such as spying or surveillance (air, land, underwater, space), defusing bombs, assisting the wounded, inspecting hideouts, and attacking targets. Police and security robots today perform similar functions, in addition to guarding borders and buildings, scanning for pedophiles and criminals, dispensing helpful information, reciting warnings, and more. There is also a growing market for home-security robots, which can shoot pepper spray or paintball pellets and transmit pictures of suspicious activities to their owners' mobile phones.

Research and education: Scientists are using robots in laboratory experiments and in the field, such as collecting ocean surface and marine-life data over extended periods (e.g., Rutgers University's Scarlet Knight) and exploring new planets (e.g., NASA's Mars Exploration Rovers). In classrooms, robots are delivering lectures, teaching subjects (e.g., foreign languages, vocabulary, and counting), checking attendance, and interacting with students.

Entertainment: Related to the above is the field of "edutainment" or education-entertainment robots, which include ASIMO, Nao, iCub, and others. Though they may lack a clear use, such as for military or manufacturing, they aid researchers in the study of cognition (both human and artificial), motion, and other areas related to the advancement of robotics. Robotic toys, such as AIBO, Pleo, and RoboSapien, also serve as discovery and entertainment platforms.

Medical and healthcare: Some toy-like robots, such as PARO which looks like a baby seal, are designed for therapeutic purposes, such as reducing stress, stimulating cognitive activity, and improving socialization. Similarly, University of Southern California's socially assistive robots help coach physical-therapy and other patients. Medical robots, such as da Vinci Surgical System and ARES ingestible robots, are assisting with or conducting difficult medical procedures on their own. RIBA, IWARD, ERNIE, and other robots perform some the functions of nurses and pharmacists.

Personal care and companions: Robots are increasingly used to care for the elderly and children, such as RI-MAN, PaPeRo, and CareBot. PALRO, QRIO, and other edutainment robots mentioned above can also provide companionship. Surprisingly, relationships of a more intimate nature are not quite satisfied by robots yet, considering the sex industry's reputation as an early adopter of new technologies. Introduced in 2010, Roxxxy is billed as "the world's first sex robot" [17], but its lack of autonomy or capacity to "think" for itself, as opposed to merely respond to sensors, suggests that it is not in fact a robot, per the definition above.

Environment: Not quite as handy as WALL-E, robots today still perform important functions in environmental remediation, such as collect trash, mop up after nuclear power plant disasters, remove asbestos, cap oil geysers, sniff out toxins, identify polluted areas, and gather data on climate warming.

In the future: As AI advances, we can expect robots to play more complex and a wider range of roles in society: For instance, police robots equipped with biometrics capabilities and sensors could detect weapons, drugs, and faces at a distance. Military robots could make attack decisions on their own; in most cases today, there is a human triggerman behind those robots. Driverless trains today and DARPA's Grand Challenges are proof-of-concepts that robotic transportation is possible, and even commercial airplanes today are controlled autonomously for a significant portion of their flight today, never mind military UAVs. A general-purpose robot, if achievable, could service many of our domestic labor needs, as opposed to a team of robots each with its own job.

We can also expect robots to scale down as well as up: Some robots are miniature today and ever shrinking, perhaps bringing to life the idea of a "nano-bot", swarms of which might work inside our bodies or in the atmosphere or cleaning up oil spills. Even rooms or entire buildings might be considered as robots—beyond the "smart homes" of today—if they can manipulate the environment in ways more significant than turning on lights and air conditioning. With synthetic biology, cognitive science, and nanoelectronics, future robots could be biologically based. And man-machine integrations, i.e., cyborgs, may be much more prevalent than they are today, which are mostly limited to patients with artificial body parts, such as limbs and joints that are controlled to some degree by robotics. Again, much of this speaks to the fuzziness of the definition of robot: What we intuitively consider as robots today may change given different form-factors and materials of tomorrow.

In some countries, robots are quite literally replacements for humans, such as Japan, where a growing elderly population and declining birthrates mean a shrinking workforce [35]. Robots are built to specifically fill that labor gap. And given the nation's storied love of technology, it is therefore unsurprising that approximately one out of 25 workers in Japan is a robot [32]. While the US currently dominates the market in military robotics, nations such as Japan and South Korea lead in the market for social robotics, such as elderly-care robots. Other nations with similar demographics, such as Italy, are expected to introduce more robotics into their societies, as a way to shore up a decreasing workforce [19]; and nations without such concerns can drive productivity, efficiency, and effectiveness to new heights with robotics.

3. Ethical and social issues

The Robotics Revolution promises a host of benefits that are compelling and imaginative, but as with other emerging technologies, they also come with risks and new questions that society must confront. This is not unexpected, given the disruptive nature of technology revolutions. In the following, we map the myriad issues into three broad (and interrelated) areas of ethical and social concern and provide representative questions for each area:

3.1. Safety and errors

We have learned by now that new technologies, first and foremost, need to be safe. Asbestos, DDT, and fen-phen are among the usual examples of technology gone wrong (e.g., [41,20,24]), having been introduced into the marketplace before sufficient health and safety testing. A similar debate is occurring with nanomaterials now (e.g., [3]).

With robotics, the safety issue is with their software and design. Computer scientists, as fallible human beings, understandably struggle to create a perfect piece of complex software: somewhere in the millions of lines of code, typically written by teams of programmers, errors and vulnerabilities likely exist. While this usually does not result in significant harm with, say, office applications—just lost data if users do not periodically save their work (which arguably is their own fault)—even a tiny software flaw in machinery, such as a car or a robot, could lead to fatal results.

For instance, in August 2010, the US military lost control of a helicopter drone during a test flight for more than 30 minutes and 23 miles, as it veered towards Washington DC, violating airspace restrictions meant to protect the White House and other governmental assets [11]. In October 2007, a semi-autonomous robotic cannon deployed by the South African army malfunctioned, killing nine "friendly" soldiers and wounding 14 others (e.g., [34]). Experts continue to worry about whether it is humanly possible to create software sophisticated enough for armed military robots to discriminate combatants from noncombatants, as well as threatening behavior from nonthreatening (e.g., [26]).

Never mind the scores of other military-robot accidents and failures [46], human deaths can and have occurred in civilian society: The first human to be killed by a robot was widely believed to be in 1979, in an auto factory accident in the US [23]. And it does not take much to imagine that a mobile city-robot—a heavy piece of machinery—could accidentally run over a small child.

Hacking is an associated concern, given how much attention is paid to computer security today. What makes a robot useful—its strength, ability to access and operate in difficult environments, expendability, and so on—could also be turned against us, either by criminals or simply mischievous persons. This issue will become more important as robots become networked and more indispensible to everyday life, as computers and smart phones are today. Indeed, the fundamentals of robotics technology are not terribly difficult to master: as formidable and fearsome as military robots are today, already more than 40 nations have developed those capabilities, including Iran [39,15].

Thus, some of the questions in this area include: Is it even possible for us to create machine intelligence that can make nuanced distinctions, such as between a gun and an ice-cream cone pointed at it, or understand human speech that is often heavily based on context? What are the tradeoffs between non-programming solutions for safety—e.g., weak actuators, soft robotic limbs or bodies, using only non-lethal weapons, or using robots in only specific situations such as a “kill box” in which all humans are presumed to be enemy targets—and the limitations they create? How safe ought robots be prior to their introduction into the marketplace or society, i.e., should a precautionary principle apply here? How would we balance the need to safeguard robots from running amok (e.g., with a kill-switch) with the need to protect it from hacking or capture? How can the “frame problem” be solved in practice, in such a way as to ensure robots only take salient, relevant safety information into account?

3.2. Law and ethics

Linked to the risk of robotic errors, it may be unclear who is responsible for any resulting harm. Product liability laws are largely untested in robotics and, anyway, continue to evolve in a direction that releases manufacturers from responsibility, e.g., end-user license agreements in software. With military robots, for instance, there is a list of characters throughout the supply chain that may be held accountable: the programmer, the manufacturer, the weapons legal review team, the military procurement officer, the field commander, the robot’s handler, and even the President of the United States, as the commander-in-chief.

As robots become more autonomous, it may be plausible to assign responsibility to the *robot itself*, e.g., if it is able to exhibit enough of the features that typically define personhood. If this seems too far-fetched, consider that synthetic biology is forcing society to reconsider the definition of life, blurring the line between living and non-living agents (e.g., [10]). Also consider that there is ongoing work in integrating computers and robotics with biological brains (e.g., [44]). A conscious human brain (and its body) presumably has human rights, and replacing parts of the brain with something else, while not impairing its function, would seem to preserve at least some of those rights and responsibilities (and possibly even add novel rights as new capacities emerge, given “ought implies can”). We may come to a point at which more than half of the brain or body is artificial, making the organism more robotic than human; seeing such a continuum between humans and robots may make the issue of robot rights and duties more plausible. And if some (future) robots or cyborgs meet the necessary requirements to have rights, which ones should they have, and how does one manage such portfolios of rights, which may be unevenly distributed given a range of biological and technological capabilities?

In the near term, one natural way to think about minimizing risk of harm from robots is to program them to obey our laws or follow a code of ethics. Of course, this is much easier said than done, since laws can be vague and context-sensitive, which robots may not be sophisticated enough to understand, at least in the foreseeable future. Even the three (or four) law of robotics in Asimov’s stories, as elegant and sufficient as they appear to be, create loopholes that result in harm (e.g., [6–8]).

Programming aside, the use of robots must also comply with law and ethics, and again those rules and norms may be unclear or untested on such issues. For instance, landmines are an effective but horrific weapon that indiscriminately kills, whether soldiers or children; they have existed for hundreds of years, but it was only in 1983—after their heavy use in 20th century wars—that certain uses of landmines were banned, e.g., planting them without means to identify and remove them later [40]; and only in 1999 did an international treaty ban the production and use of landmines [1]. Likewise, the use of military robots may raise legal and ethical questions that we have yet to fully consider (e.g., [26,27]) and, later in retrospect, may seem obviously unethical or unlawful.

Another relevant area of law concerns privacy. Several forces are driving this concern, including: the shrinking size of digital cameras and other recording devices, an increasing emphasis on security at the expense of privacy (e.g., expanded wiretap laws, a blanket of surveillance cameras in some cities to monitor and prevent crimes), advancing biometrics capabilities and sensors, and database integrations. Besides robotic spy planes, we previously mentioned (future) police robots that could conduct intimate surveillance at a distance, such as detecting hidden drugs or weapons and identifying faces unobtrusively; if linked to databases, they could also run background checks on one’s driving, medical, banking, shopping, or other records to determine if the person should be apprehended [36]. Domestic robots too can be easily equipped with surveillance devices—as home security robots already are—that may be monitored or accessed by third parties [12].

Of course, ethical and cultural norms, and therefore law, vary around the world, so it is unclear *whose* ethics and law ought to be the standard in robotics; and if there is no standard, which jurisdictions would gain advantages or cause policy challenges internationally? Such challenges may require international policies, treaties, and perhaps even international laws and enforcement bodies. This kind of political-cultural schism is not merely theoretical, but one we have already seen in military law: the US, for instance, has refused to sign or accede the aforementioned landmine ban, also known as the Ottawa Treaty. The relationship of robotic aircraft (drones) to international law is likewise a vexed issue: the US assumes its Predator attacks in Pakistan are legal, whereas many other countries disagree. And Japanese and Americans may well have different moral sensibilities in leaving care of the elderly solely in the hands (or extender arms) of robots. From these global variations in ethics and law, it may be reasonable to expect an uneven trajectory for robot development worldwide, which affects the proliferation of associated benefits and pragmatic challenges.

Other questions in this area include: If we could program a code of ethics to regulate robotic behavior, which ethical theory should we use? Are there unique legal or moral hazards in designing machines that can autonomously kill people? Or should robots merely be considered tools, such as guns and computers, and regulated accordingly? Is it ethically permissible to abrogate responsibility for our elderly and children to machines that seem to be a poor substitute for human companionship (but perhaps better than no—or abusive—companionship)? Will robotic companionship (that could replace human or animal companionship) for other purposes, such as drinking buddies, pets, other forms of entertainment, or sex, be morally problematic? At what point should we consider a robot to be a "person", thus affording it some rights and responsibilities, and if that point is reached, will we need to emancipate our robot "slaves"? Do we have any other distinctive moral duties towards robots? As they develop enhanced capacities, should cyborgs have a different legal status than ordinary humans? At what point does a technology-mediated surveillance count as a "search", which would generally require a judicial warrant? Are there particular moral qualms with placing robots in positions of authority, such as police, prison or security guards, teachers, or any other government roles or offices in which humans would be expected to obey robots?

3.3. Social impact

How might society change with the Robotics Revolution? As with the Industrial and Internet Revolutions, one key concern is about a loss of jobs. Factories had replaced legions of workers who used to perform the same work by hand, giving way to the faster, more efficient processes of automation. Internet ventures such as Amazon.com, eBay, and even smaller "e-tailers" are still edging out brick-and-mortar retailers who have much higher overhead and operating expenses. Likewise, as potential replacements for humans—performing certain jobs better, faster, and so on—robots may displace human jobs, regardless of whether the workforce is growing or declining.

The standard response is that human workers, whether replaced by other humans or machines, would then be free to focus their energies where they can make a greater impact, i.e., at jobs in which they have a greater competitive advantage [33]; to resist this change is to support inefficiency. For instance, by outsourcing call-center jobs to other nations where the pay is less, displaced workers (in theory) can perform "higher-value" jobs, whatever those may be. Further, the demand for robots itself creates additional jobs. Yet, theory and efficiency provide little consolation for the human worker who needs a job to feed her or his family, and cost-benefits may be negated by unintended effects, e.g., a negative customer experience with call-center representatives whose first language is not that of the customers.

Connected to labor, some experts are concerned about technology dependency (e.g., [42]). For example, as robots prove themselves to be better than humans in difficult surgeries, the resulting loss of those jobs may also mean the gradual loss of that medical skill or knowledge, to the extent that there would be fewer human practitioners. This is not the same worry with labor and service robots that perform dull and dirty tasks, in that we care less about the loss of those skills; but there is a similar issue of becoming overly reliant on technology for basic work. For one thing, this dependency seems to cause society to be more fragile: for instance, the Y2K problem caused significant panic, since so many critical systems—such as air-traffic control and banking—were dependent on computers whose ability to correctly advance their internal clock to January 1, 2000 (as opposed to resetting it to January 1, 1900) was uncertain; and similar situations exist today with malicious computer *viruses du jour*.

Like the social networking and email capabilities of the Internet Revolution, robotics may profoundly impact human relationships. Already, robots are taking care of our elderly and children, though there are not many studies on the effects of such care, especially in the long term. Some soldiers have emotionally bonded with the bomb-disposing PackBots that have saved their lives, sobbing when the robot meets its end (e.g., [38,22]). And robots are predicted to soon become our lovers and companions [25]: they will always listen and never cheat on us. Given the lack of research studies in these areas, it is unclear whether psychological harm might arise from replacing human relationships with robotic ones.

Harm also need not be directly to persons, e.g., it could also be to the environment. In the computer industry, "e-waste" is a growing and urgent problem (e.g., [31]), given the disposal of heavy metals and toxic materials in the devices at the end of their product lifecycle. Robots as embodied computers will likely exacerbate the problem, as well as increase pressure on rare-earth elements needed today to build computing devices and energy resources needed to power them. Networked robots would also increase the amount of ambient radiofrequency radiation, like that created by mobile phones—which have been blamed, fairly or not, for a decline of honeybees necessary for pollination and agriculture [37], in addition to human health problems (e.g., [2]).

Thus, some of the questions in this area include: What is the predicted economic impact of robotics, all things considered? How do we estimate the expected costs and benefits? Are some jobs too important or too dangerous for machines to take over? What do we do with the workers displaced by robots? How do we mitigate disruption to a society dependent on robotics, if those robots were to become inoperable or corrupted, e.g., through an electromagnetic pulse or network virus? Is there a danger with emotional attachments to robots? Are we engaging in deception by creating anthropomorphized machines that may lead to such attachments, and is that bad? Is there anything essential in human companionship and relationships that robots cannot replace? What is the environmental impact of a much larger robotics industry than we have today? Could we possibly face any truly cataclysmic consequences from the widespread adoption of social robotics, and if so, should a precautionary principle apply?

4. Engaging the issues now

These are only some of the questions which the emerging field of robot ethics is concerned with, and many of these questions lead to the doorsteps of other areas of ethics and philosophy, e.g., computer ethics and philosophy of mind, in addition to the disciplines of psychology, sociology, economics, politics, and more. Note also that we have not even considered the more popular "Terminator" scenarios in which robots—through super artificial intelligence—subjugate humanity, which are highly speculative scenarios that continually overshadow more urgent and plausible issues.

The robotics industry is rapidly advancing, and robots in society today are already raising many of these questions. This points to the need to attend to robot ethics now, particularly as consensus on ethical issues is usually slow to catch up with technology, which can lead to a "policy vacuum" [30]. As an example, the Human Genome Project was started in 1990, but it took 18 years after that for Congress to finally pass a bill to protect Americans from discrimination based on their genetic information. Right now, society is still fumbling through privacy, copyright, and other intellectual property issues in the Digital Age, nearly 10 years since Napster was first shut down.

As researchers and educators, we hope that our forthcoming volume of robot-ethics papers will provide and motivate greater discussion—both in and outside the classroom—across the broad continuum of issues, such as described above. The contributors to our edited book include many respected and well-known experts in robotics and technology ethics today, including: Colin Allen, Peter Asaro, Anthony Beavers, Selmer Bringsjord, Marcello Guarini, James Hughes, Gert-Jan Lokhorst, Matthias Scheutz, Noel Sharkey, Rob Sparrow, Jeroen van den Hoven, Gianmarco Veruggio, Wendell Wallach, Kevin Warwick, and others.

Sometimes to deaf ears, history lectures us on the importance of foresight: While the invention of such things as the printing press, gunpowder, automobiles, computers, vaccines, and so on, has profoundly changed the world (for the better, we hope), they have also led to unforeseen consequences, or perhaps consequences that might have been foreseen and addressed had we bothered to investigate them. Least of all, they have disrupted the *status quo*, which is not necessarily a terrible thing in and of itself; but unnecessary and dramatic disruptions, such as mass displacements of workers or industries, have real human costs to them. Given lessons from the past, society is beginning to think more about ethics and policy in advance of, or at least in parallel to, the development of new game-changing technologies, such as genetically-modified foods, nanotechnology, neuroscience, and human enhancement—and now we add robotics to that syllabus.

At the same time, we recognize that these technologies seem to jump out of the pages of science fiction, and the ethical dilemmas they raise also seem too distant to consider, if not altogether unreal. But as Isaac Asimov foretold: "It is change, continuing change, inevitable change, that is the dominant factor in society today. No sensible decision can be made any longer without taking into account not only the world as it is, but the world as it will be ... This, in turn, means that our statesmen, our businessmen, our everyman must take on a science fictional way of thinking" [7]. With human ingenuity, what was once fiction is becoming fact, and the new challenges it brings are all too real.

Acknowledgements

Some of this paper is excerpted from our book *Robot Ethics: The Ethical and Social Implications of Robotics*, to be published by MIT Press in late 2011 [28]. We would like to thank Kyle Campbell for his editorial assistance with this paper.

References

[1] Jeff Abramson, Arms Control Association: The Ottawa Convention at a Glance, June 2008. Accessible at http://www.armscontrol.org/factsheets/ottawa. Last accessed on September 12, 2010.
[2] Ashok Agarwal, Nisarg R. Desai, Kartikeya Makker, Alex Varghese, Rand Mouradi, Edmund Sabanegh, Rakesh Sharma, Effects of radiofrequency electromagnetic waves (RF-EMW) from cellular phones on human ejaculated semen: an in vitro pilot study, Fertility and Sterility 92 (4) (2009) 1318–1325.
[3] Fritz Allhoff, Patrick Lin, Daniel Moore, What Is Nanotechnology and Why Does It Matter?: From Science to Ethics, Wiley-Blackwell, Hoboken, NJ, 2010.
[4] Ronald C. Arkin, Governing lethal behavior: embedding ethics in a hybrid deliberative/hybrid robot architecture, Report GIT-GVU-07-11, Georgia Institute of Technology's GVU Center, Atlanta, GA, 2007. Accessible at http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf. Last accessed on September 12, 2010.
[5] Isaac Asimov, I, Robot, 2004 edition, Bantam Dell, New York, NY, 1950.
[6] Isaac Asimov, The Naked Sun, Doubleday, New York, NY, 1957.
[7] Isaac Asimov, My own view, in: Robert Holdstock (Ed.), The Encyclopedia of Science Fiction, St. Martin's Press, New York, NY, 1978.
[8] Isaac Asimov, Robots and Empire, Doubleday, New York, NY, 1985.
[9] George Bekey, Autonomous Robots: From Biological Inspiration to Implementation and Control, MIT Press, Cambridge, MA, 2005.
[10] Steven A. Benner, Q&A: life, synthetic biology, and risk, BMC Biology 8 (2010) 77.
[11] Elisabeth Bumiller, The New York Times: Navy Drone Violated Washington Airspace (Aug. 25, 2010).
[12] M. Ryan Calo, Robots and privacy, in: Patrick Lin, George Bekey, Keith Abney (Eds.) Robot Ethics: The Ethical and Social Implications of Robotics, MIT Press, Cambridge, MA, 2010 (forthcoming). Accessible at http://ssrn.com/abstract=1599189. Last accessed on September 12, 2010.
[13] Karel Čapek, Rossum's Universal Robots, 2004 edition, Penguin Group, New York, NY, 1921 (trans. Claudia Novack).
[14] Keay Davidson, How a butterfly's wing can bring down Goliath, San Francisco Chronicle (August 15, 2003).
[15] Defense Update, Karrar–Iran's new jet-powered recce and attack drone, 2010. Accessible at http://defense-update.com/products/k/karrar_jet_powered_drone_24082010.html. Last accessed on September 12, 2010.
[16] Philip K. Dick, Do Androids Dream of Electric Sheep?, Del Rey Books, New York, NY, 1968.
[17] Yvonne Fulbright, Meet Roxxxy, the 'woman' of your dreams, 2010. Accessible at http://www.foxnews.com/story/0,2933,583314,00.html. Last accessed on September 12, 2010.

[18] Bill Gates, A robot in every home, Scientific American (January 2007) 58–65.
[19] Gary Geipel, Global aging and the global workforce, a Hudson Institute article, 2003. Accessible at http://www.hudson.org/index.cfm?fuseaction=publication_details&id=2740. Last accessed on September 12, 2010.
[20] Christine Gorman, Danger in the diet pills?, Time Magazine (July 21, 1997). Accessible at http://www.time.com/time/magazine/article/0,9171,986725,00.html. Last accessed on September 12, 2010.
[21] Erico Guizzo, IEEE Spectrum: World Robot Population Reaches 8.6 Million, April 14, 2010. Accessible at http://spectrum.ieee.org/automaton/robotics/industrial-robots/041410-world-robot-population. Last accessed on September 12, 2010.
[22] Jeremy Hsu, Real soldiers love their robot brethren, Live Science (May 21, 2009). Accessible at http://www.livescience.com/technology/090521-terminator-war.html. Last accessed on September 12, 2010.
[23] Tim Kiska, Death on the job: Jury awards $10 million to heirs of man killed by robot at auto plant, Philadelphia Inquirer (August 11, 1983) A-10.
[24] Linda Lear, Rachel Carson: Witness for Nature, Henry Hoyten, New York, NY, 1997.
[25] David Levy, Love and Sex with Robots: The Evolution of Human–Robot Relationships, Harper Collins Publishers, New York, NY, 2007.
[26] Patrick Lin, George Bekey, Keith Abney, Autonomous military robots: risk, ethics, and design. A report commissioned by US Department of Navy/Office of Naval Research, 2008. Accessible at http://ethics.calpoly.edu/ONR_report.pdf. Last accessed on September 12, 2010.
[27] Patrick Lin, George Bekey, Keith Abney, Robots in war: issues of risk and ethics, in: Rafael Capurro, Michael Nagenborg (Eds.), Ethics and Robotics, AKA Verlag/IOS Press, Heidelberg, Germany, 2009.
[28] Patrick Lin, Keith Abney, George Bekey, Robot Ethics: The Ethical and Social Implications of Robotics, MIT Press, Cambridge, MA (forthcoming).
[29] Alexis Madrigal, Market data firm spots the tracks of bizarre robot traders, The Atlantic (August 4, 2010).
[30] James H. Moor, Metaphilosophy: What is computer ethics? 16 (4) (1985) 266–275.
[31] Amy Joi O'Donoghue, E-waste is a growing issue for states, Deseret News (August 22, 2010). Accessible at http://www.deseretnews.com/article/700059360/E-waste-is-a-growing-issue-for-states.html?pg=1. Last accessed on September 12, 2010.
[32] RedOrbit, Japan hopes to employ robots by 2025, April 8, 2008. Accessible at http://www.redorbit.com/news/technology/1332274/japan_hopes_to_employ_robots_by_2025/. Last accessed on September 12, 2010.
[33] Mitch Rosenberg, The surprising benefits of robots in the DC, Supply & Demand Chain Executive (June/July 2009).
[34] Noah Schactman, Robot cannon kills 9, wounds 14, Wired (October 18, 2007). Accessible at http://blog.wired.com/defense/2007/10/robot-cannon-ki.html. Last accessed on September 12, 2010.
[35] Chana Schoenberger, Japan's shrinking workforce, Forbes (May 25, 2008). Accessible at http://www.forbes.com/2008/05/25/immigration-labor-visa-oped-cx_crs_outsourcing08_0529japan.html. Last accessed on September 12, 2010.
[36] Noel Sharkey, 2084: big robot is watching you, A Commissioned Report, 2008. Accessible at http://www.dcs.shef.ac.uk/~noel/. Last accessed on September 12, 2010.
[37] Ved Parkash Sharma, Neelima R. Kumar, Changes in honeybee behaviour and biology under the influence of cellphone radiations, Current Science 98 (10) (2010) 1376–1378.
[38] Peter W. Singer, Wired for War: The Robotics Revolution and Conflict in the 21st Century, Penguin Press, New York, NY, 2009.
[39] Peter W. Singer, Robots at war: the new battlefield, Wilson Quarterly (Winter 2009). Accessible at http://www.wilsonquarterly.com/article.cfm?aid=1313. Last accessed on September 12, 2010.
[40] United Nations, The Convention on Certain Conventional Weapons, Entered into force on December 2, 1983. Accessible at http://www.armscontrol.org/factsheets/CCW. Last accessed on September 12, 2010.
[41] US Environmental Protection Agency, Asbestos ban and phase out, April 25, 2007. Accessible at http://www.epa.gov/asbestos/pubs/ban.html. Last accessed on September 12, 2010.
[42] Gianmarco Veruggio (Ed.), EURON Roboethics Roadmap, EURON Roboethics Atelier, EURON, Genoa, Italy, 2006. Accessible at http://www.roboethics.org/atelier2006/docs/ROBOETHICS%20ROADMAP%20Rel2.1.1.pdf. Last accessed on September 12, 2010.
[43] Wendell Wallach, Colin Allen, Moral Machines: Teaching Robots Right From Wrong, Oxford University Press, New York, 2009.
[44] Kevin Warwick, Implications and consequences of robots with biological brains, in: Anthony Beavers (Ed.), Special Issue on Robot Ethics and Human Ethics, Ethics & Information Technology 12 (1) (2010).
[45] Daniel H. Wilson, How to Survive a Robot Uprising: Tips on Defending Yourself Against the Coming Rebellion, Bloomsbury Publishing, New York, NY, 2005.
[46] David Zucchino, War zone drone crashes add up, Los Angeles Times (July 6, 2010).