HOW TO GROW A ROBOT

Developing Human-Friendly, Social AI

Mark Lee

This book was set in Stone Serif by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Lee, Mark H., author.
Title: How to grow a robot : developing human-friendly, social AI / Mark Lee.
Description: Cambridge : The MIT Press, 2020. | Includes bibliographical references and index.
Identifiers: LCCN 2019025740 | ISBN 9780262043731 (hardcover) | ISBN 9780262357852 (ebook)
Subjects: LCSH: Robots—Control systems. | Robots—Social aspects. | Machine learning. | Artificial intelligence—Forecasting. |
Human-computer interaction.
Classification: LCC TJ211.35 .L43 2020 | DDC 006.301—dc23
LC record available at https://lccn.loc.gov/2019025740

CONTENTS

PREFACE

I have always been fascinated by the relationship between machines and humans. In my youth, I was very interested in electrical machines and engineering generally, which, with hindsight, I now see as an incredibly creative discipline. But during my training as an engineer, I was particularly attracted to systems that appeared to mimic some aspects of human behavior. Even simple feedback systems can be captivating (e.g., watching a ship constantly correct its course while being buffeted about by winds and tides). This leads to questions like: What kind of mental machinery do humans actually use? I somehow managed to work on projects that combined engineering with human-centered issues: speech encoding, color vision processing, and autonomous control. I found that psychology was the vital missing element, and my PhD, in modeling sensorimotor control and coordination, was a precursor of the work described here. The relationships among computers, brains, and machines are many and fascinating.

When I tried to pursue this line of research, in the 1980s, it was completely out of favor. The new commercialism of artificial intelligence (AI) had just begun, and AI software was becoming highly marketable. Before then, the early pioneers of AI saw no large distinctions between human and computer intelligence and thought both could be studied together. The drive for software products and applications caused AI to become largely separate from the study of human cognition. I remember discussing my work with Andy Barto, Richard Sutton, and others at the University of Massachusetts–Amherst when they were in the middle of developing reinforcement learning. In the UK, it was very difficult to get any funding for basic, curiosity-driven AI research, so I started work on AI in industrial robotics, tactile sensing, and error recovery and diagnostic systems, still exploring how humans behave during such tasks and how AI could be used in robotics. Fortunately, the twenty-first century has seen a swing back to topics like learning, autonomous agents, and renewed attention to human performance, and there is now a strong global research community in the new field of developmental robotics.

I have written this book partly in response to some common misconceptions about robots and AI. A great deal of misinformation about robots and AI has been bandied about in the media, some of which is clearly unreasoned nonsense. While there will be some amazing technology produced in the near future, there are limitations inherent in AI, and there are ethical issues that involve us all.

However, this is not a negative story because, after explaining the difficulties and nature of the problem, I then go on to describe an alternative approach that is currently delivering results and showing real potential for developing general-purpose robots. I give details from my own and related research to show how psychology and developmental ideas can get nearer to humanlike behavior—which is fundamentally general purpose but can adapt in order to learn to accomplish specialized tasks.

I try to avoid the minutiae of the latest technology or subtle changes of fashion; instead, I concentrate on the trends, the way technology develops, and the role of technology in our lives. I feel that it is important that everyone should be much more engaged in appreciating the technological developments of the day in order to get a better perspective on what is feasible, what is reasonable, and what is wild hyperbole. By applying basic common sense to many of the claims and predictions, we can evaluate, and therefore have a much larger influence on, the role of technology in our lives (and future lives). This book is a response to these concerns and contains some of my findings and insights into these fascinating problems. I hope that it will give a perspective that offers a better way of approaching some of the wider issues raised by intelligent robots.

ACKNOWLEDGMENTS

There are many colleagues, friends, and students that have helped me in my career, and it is impracticable to list them all. This includes colleagues in the Computer Science Department at Aberystwyth University, as well as many others in universities and other organizations in the UK, Europe, Japan, and the US. Research is a truly global community. I am grateful to all these friends and colleagues who generously supported, influenced, and encouraged the approach described here. I should particularly thank those who were directly involved in the work reported in this book: Raphael Braud, Fei Chao, Kevin Earland, Tao Geng, Alexandros Giagkos, Richard Gunstone, Martin Hülse, Ian and Tom Izzett, Suresh Kumar, James Law, Daniel Lewkowicz, Sebastian McBride, Qinggang Meng, Marcos Rodrigues, Patricia Shaw, Michael Sheldon, Qiang Shen, James Wilson, and Xing Zhang. Of course, the opinions expressed here are solely my responsibility, as are the errors and omissions.

Others who have helped in more mysterious ways include Anna Borghi, Merideth Gattis, Frank Guerin, Kevin Gurney, David Llewelyn, Giorgio Metta, David Middleton, Kevin O'Regan, Peter Redgrave, Peter Tallack, Raymond Tallis, and my late but dear friends Brendan McGonigle and Ulrich Nehmzow. Forgive me for not listing all the others; you are not forgotten, and I am grateful to you all.

I much appreciate the expertise and kind support of Marie L. Lee at MIT Press, and I thank all the copy-editing and other staff who have helped in this project.

The research described here was mainly carried out during four projects: two funded by the British Engineering and Physical Science Research Council, and two funded by the FP7 program by the European Commission (EC). The iCub humanoid robot originated in an EC-funded project, and the EC had the foresight to support the supply of these robots to new research projects, with the result that iCubs are now being used in more than twenty-five robotics research laboratories worldwide.

I am grateful to the Royal Society of Arts for permission to reuse and modify material taken from my article "A Frame of Mind," *Royal Society of the Arts Journal*, no. 2, 2018, 46–47. This appears in chapter 9, in the section entitled "The Brain Is a Machine—So What?"

I am grateful for the use of many libraries, particularly six local libraries. This includes the Bronglais Postgraduate Center and the wonderful resources of Llyfrgell Genedlaethol Cymru, the National Library of Wales. Finally, I am most grateful for having such a positive, helpful, and brilliant family, and particularly for my dear wife, Elizabeth, who has encouraged and supported this project in so many ways.

I

WHAT'S WRONG WITH ARTIFICIAL INTELLIGENCE?

1.1 Keepon displaying eye contact and joint attention (looking at an object that the human has just looked at). Used with permission from Professor Hideki Kozima (Tohoku University).

creatures move. Walt Disney created an industry, and made a fortune, from this effect.

Keepon is captivating because it *looks at humans when they talk* and also performs what's known as *joint attention*—it looks at the same objects that the human looks at, giving the impression of shared interest (see figure 1.1). It's not necessary for a social robot to pretend to be human, or even resemble us. We are happy talking to all kinds of animals: Just look at the fuss we make over a puppy or a kitten! But we do need that animal-like behavior and response.

This raises a big question: If Keepon and some other robots can give a passable imitation of animal movements, why do the vast majority of robots seem so clunky, so "mechanical"? If we want future robots to become useful assistants, companions, or supernumeraries, then they will have to be much more humanlike, friendly, and engaging.

ACTING AND THINKING

Robotic devices have fascinated both scientists and the public at large for a very long time. Some of the earliest ones were treated as religious marvels, such as the automaton monk of 1560, a small figurine that prayed while "walking" across a table top.[2] From the seventeenth century onward, automata became very popular as entertainments, and the eighteenth century was a golden age for autonomous mechanical figures and animals. Thanks to the skills of watchmakers, really impressive models were constructed, such as the life-size Silver Swan of 1773, which moves its head and neck realistically, preens itself, and catches a wriggling fish from the simulated moving water on which it sits.[3] The automatic human figures of those times often played musical instruments or danced; sometimes they wrote or drew on paper; but they always mimicked human behavior. The key to their attraction was that their actions were humanlike.

From the clockwork toys and automata that so amused our ancestors to the amazing science fiction creations seen in modern film and television, we have always had a deep empathy for machines that can mimic human actions. This fascination extends across a wide range of movement styles, from total mimicry to imitations of bodily actions. We enjoy watching the grace and elegance of the human form in athletics or ballet, but we are also intrigued by the motions of factory machines that assemble components or paint cars ("poetry in motion," as I saw emblazoned on the side of a heavy lifting crane).

It's not just movement: We are also fascinated by displays of purely mental activity. Champion chess players, quiz masters, and expert players at word and number games all draw large audiences; and thinkers like Isaac Newton, Charles Darwin, and Albert Einstein are revered as scientific heroes. Future robots will have to be smart too—very smart. If they are going to keep up with humans, then they will have to understand us and what we want, and this requires a lot of brain power. They will have to perceive us, as humans, and continually learn about our changing wants and needs. And all this must happen in real time; that is, *in time with us*—not too fast or too slow. We soon ditch technology when it gets boring, inflexible, or unresponsive.

These two requirements, *appropriate behavior* and *mental ability*, are the subject of this book. Some robots are good at moving; some computers are good at thinking; but both skills are necessary for really intelligent robots. We need robots that humans can engage with, robots that are perceptive, animated, and responsive to us. This book explores if and how this can be achieved; and what is necessary to support such an advance. In short, the question is: *How can we build intelligent robots that think and behave like humans?*

THE SOCIAL ROBOT

Among the robots that already exist, very few even attempt to behave like humans. Keepon is in a very small minority. Most robots can be called *autonomous agents*, but very few are *social agents*. This is the issue that this book examines: Robots that think and behave like humans must be social agents, freely interacting with humans and accepted by humans.

Just to be clear, a social robot is more than just a fancy interface, avatar, or conversational system. Social robots act like a companion or friend, with frequent interactions over an extended period. They won't necessarily look like, or pretend to be human, but they will understand human gestures, behaviors, and communications and be able to respond in ways that we find socially satisfactory. They will have limited "emotional" behavior but they will learn and adapt through their social life, recognizing different people and remembering their personalities from previous encounters. Of course, they will also do useful work, but rather than being programmed for specific tasks, they will be trained by their users and learn through demonstration and human guidance, as we do. If we want to create long-term, useful robot assistants, then social interaction is the key.

We will see in chapter 8 how competence in certain kinds of dialogue has been demonstrated. But this is a narrow application: The big question is how far we can progress toward sustained and meaningful social intercourse.

This book is not concerned with any specific application of social robotics; the social robot is proposed here as a milestone for the highest level of robotic achievement: thinking and behaving like humans. This is a research goal, not an application target. For example, I don't advocate humanoid robots for the care of elderly people; this important area still requires the human touch.

But once we know how to create social robots, once we have developed the technology, there will be plenty of scope for applications, practically anywhere that human-robot interactions regularly occur.

THE ROLE OF ARTIFICIAL INTELLIGENCE

So, how can we build social robots? What techniques are required, and what knowledge do we need? The growth of artificial intelligence (AI) is widely seen as the answer. AI is viewed as the bedrock on which future robots will be created. (I use a simple definition of the term here: *AI* refers to computer systems that achieve results that are assumed to require intelligence.)

It's true that AI is making enormous strides, producing impressive advances and products. For example, IBM's Watson system can answer questions posed in everyday language. Watson dissects the question and assembles answers using huge data sources: over 200 million pages of data, including the full text of Wikipedia! This quiz machine, which won $1 million on the *Jeopardy!* game show playing against two highly expert contestants, is now being adapted for use in health care, finance, telecommunications, and biotechnology.

In other areas, like speech and image analysis, new deep learning techniques have advanced AI. Computer recognition of images now demonstrates error rates of less than 1 percent. Image processing, like analyzing X-ray images or labeling the content of pictures, is producing impressive applications and products. This new boom in AI is generating a great deal of investment. Small AI companies are frequently bought for hundreds of millions of dollars. The heads of big companies believe that they need an AI department; and often go out and buy one—or several! Robotics has always involved AI, and many future robot systems will be powered by these new AI methods.

INTELLIGENCE IN GENERAL

There are three broad strands to human intelligence: the intrinsic abilities and adaptive potential of the individual brain, the joint results of combining several brains in social interactions, and the collective cultural knowledge produced by and available to humanity (see box 1.1). AI has mostly concentrated on building versions of mental skills and competences found in the brain: an *individual* brain. This ignores the interactions *between* brains that form a crucial context for human behavior

Box 1.1

Sources of intelligence

Brains The origins and expression of intelligence are found in animal brains. This obvious fact has led to most AI systems being constructed as a single agent, thus reinforcing the idea of the individual as the sole source of intelligence.

Social interactions Any activity that requires a group of individuals to complete, such as problem-solving in emergency situations, will involve social cooperation and organization. Social conversation involves communication, shared understanding of events and experiences, and learning through language.

Culture The term *culture*, in this context, refers to the cumulative collection of all written text, data, diagrams, reference materials, and other works that humanity has compiled. Through reuse and training, such knowledge enables individuals to become more intelligent and also, through study, to discover further knowledge.

(box 1.1). In our early years, our comprehension, worldly understanding, knowledge, skills, and other cognitive abilities all depended on the interactions of at least one attendant caregiver. Without parental care, we do not develop and thrive and often do not even survive. In adult life, we form groups, societies, organizations, and networks, which are important for our continued development. Experts consult their peers and predecessors, thereby increasing their intellectual skill and knowledge. The third aspect of intelligence is culture. By *culture,* I do not mean trends, fashions, or popular activities, but rather the totality of recorded human experience across the global population. Language gives humans these extensions of intelligence into social and cultural dimensions; this is perhaps our only real difference from other animals, as will be discussed in later chapters.

Many AI systems are individual packages of intellectual skills, such as game-playing programs. On the other hand, speech and textual processing systems (e.g., translation) rely on enormous data banks of externalized knowledge. Cumulative cultural knowledge is recorded and disseminated in all kinds of ways; recently, this has become important for AI and is often known as *Big Data*—large data repositories that capture vast quantities of text, images, and other information.

We must consider both robotic hardware and pure thinking machines because, as we will see later, behaving and thinking are intimately interrelated in humans and animals, and probably need to be so in robots too. You have to be able to think *and* move the pieces to play a game like chess;[4] but who would have thought that physically moving the pieces would turn out to be the harder task to automate?

PART I: THE LIMITATIONS OF AI FOR ROBOTICS

The first part of this book looks into the state of current robotics and AI. We begin in the next chapter, with a brief overview of a range of commercial robot offerings that is presented in order for readers to appreciate the scope of modern robotic devices and what they do. This gives a picture of their current capabilities, achievements, and limitations. This is followed in chapter 3 by an overview of research robotics, including the way that innovation is supported and funded. A summary table shows two key factors that determine the difficulty of implementing successful robots: the nature of the tasks that they are set to do, and the complexity and chaos in the environments in which they have to work.

Because AI is key to future robotics, part I contains several chapters looking at how AI actually works, what it can deliver, and what it has trouble doing. Each chapter introduces a key topic that readers can use to build an understanding of the ideas and issues in modern AI and their relevance for robotics.

It turns out that there are some serious obstacles when it comes to humanlike behavior. Despite media claims, AI is nothing like human intelligence and has not made much progress in moving in this direction. AI is very powerful for specific tasks and excellent for solving individual problems, but it still struggles to reproduce the general-purpose intelligence characteristic of the human species. No program has yet been written that can pretend to be a human in any convincing way for more than a few minutes. The AI chapters lead to a striking conclusion: *Engineering methods and AI are not sufficient (or even appropriate) for creating humanlike robots*.

Humans are *generalists*, unlike many animals. We readily adapt to fit different niches; we even adapt niches to suit us. This requires a wide-ranging ability to adapt to new situations. Humans are autonomous, motivated by internal drives as much as external forces, and the key to our flexibility is our open-ended learning ability. This allows us to face new challenges, learn about them, and find ways of dealing with them. Some animals are prewired by their genes; ducklings will follow the first "parent" they see after hatching. But humans are able to learn how to master completely new tasks from scratch. Even our failures teach us important lessons. So a humanlike robot really requires general-purpose intelligence that has the potential to deal with any task.

General-purpose intelligence is the "holy grail" of AI, but thus far it has proved elusive. Despite sixty years of research and theorizing, general-purpose AI has not made any progress worth mentioning, and this could even turn out to be an impossible problem (at least, impossible for software-based AI). However, even if general-purpose AI could somehow be realized in a computer (which I doubt), this would not be enough to solve the problem of thinking and behaving like humans. This is because human intelligence is *embodied*: All our thoughts, sensations, intentions, relationships, concepts, and meanings only make sense in terms of a body. Our brains do not run abstract codes that control our behavior—that's the wrong metaphor. Brains (i.e., our thoughts) are deeply dependent on, and shaped by, our bodily experience, including social and cultural experience. We are so fully integrated into our living environments that our internal thinking structures mirror and match these in an almost symbiotic way.

The AI chapters build up from basic concepts toward recent advances in Big Data and Deep Learning (chapter 7). These developments have made some remarkable progress and are being heralded as the solution to all remaining AI problems. The idea is that AI learning methods are now (or soon will be) powerful enough to learn any task required, and thus evolve into general-purpose intelligence. This would be a serious and important event in the history of AI, and so it has profound implications. Toward the end of part I, chapters 9 and 10, on brain models and AI integration, respectively, explain the issues with general AI and the nature of agency and embodiment. These two chapters contain the core of the argument against AI in social robots.

It is possible to skip some of the AI details, but do be sure to read the list of short points ("Observations") at the end of each chapter. These state the bottom lines that summarize all the important things to remember, and they cumulatively build up the book's arguments.

PART II: DEVELOPMENTAL ROBOTICS

The obvious question then is: If AI faces obstacles in building humanlike robots, are there other ways to go? Are there any other paradigms we can explore? The second part of this book presents a very different approach. This new viewpoint says: If you want a humanlike robot, why not start from what we know about humans? The idea here is not to "build your own robot," but rather to "grow your own robot." Don't load up your robot with piles of preprepared human knowledge; allow it to explore for itself and build its own knowledge about the things that come to have meaning for it, as we do. This is an embodied, enactive approach that allows cognitive functions to grow in step with experience. The idea is to shift the source of knowledge and meaning from the robot designer, or from Big Data, to the robot itself. Inspiration and data for this approach come from studies of early infant behavior, where many new and striking ideas on learning and interactive behavior can be found.

Infants have very limited perceptual and motor skills at birth but rapidly become physically and socially adept. They learn to communicate, they build an understanding of themselves as a "self," and they learn to recognize other social agents. By modeling some of these mental growth processes, robots will be able to *develop* and learn social interaction abilities, rather than being given social algorithms by their designers or extracting behavioral patterns from human data files.

This part of the book is supported by my own research into development as a robot learning mechanism. Experiments on the iCub humanoid robot have refined the approach and produced interesting results. The robot has developed from very poor levels of sensory-motor competence (just a few reflex actions and weak sensory resolution) to gaining mastery of the muscles in its limbs, eyes, head, and torso, perceiving and reaching for interesting stimuli, and playing with objects. These longitudinal experiments have taken the robot from equivalent infant stages of newborn to about nine months. At present, the iCub can reach and play with objects, associate single word sounds with familiar objects or colors, make meaningful gestures, and produce new actions for new situations.

Fortunately, the way that babies play and interact is well documented in the large body of theoretical and experimental work found in developmental psychology. This provides both inspiration and fiducial milestones for the new research field of developmental robotics. The origins of this kind of robotics are explained in chapter 11, as is embodiment, which is fundamental and different from AI. Then, the way that the iCub works is described and the principles made clear. It becomes obvious that robots can have a kind of subjective experience because everything learned is based on active experience *as seen by the robot*. By developing learning, they generalize and refine their most basic experiences so that meaningful patterns and correlations are stored, while unreliable phenomena are discarded. Note that these robots develop individually; they may begin identical, but they will produce somewhat different mental models and behaviors. This reminds us of the uniqueness of humans: Each person has her or his own history of lifetime experiences, and these contextualize particular social interactions.

This "growing" methodology, involving repeated refinements through experimentation, is synthetic rather than analytic and, I believe, gives us a much better route toward human-friendly robots with more humanlike characteristics.

PART III: FURTHER AHEAD, LIKELY DEVELOPMENTS, AND FUTUROLOGY

The implications for the future of both AI robots and humanlike robots are outlined in the final part of this book. The likely near-term advances in behavior and performance are estimated and compared. For instance, AI robots are given goals by their designer; they have an objective stance because they are structured from without. So the contrast with developmental robots is sharp and revealing. For example, future AI robots will have vast data resources available to them. They will be able to answer all kinds of questions, they will solve problems, and they will learn. So entering an unknown room, recognizing objects, and carrying out tasks using the objects in the room will become quite routine. Their abilities with objects and environments will initially be much superior to developmental robots. But their behavior will not really be humanlike. They will use Big Data on human behavior to simulate common human patterns, often convincingly. But they will solve problems differently, reason differently, and construct different concepts and percepts. And these differences will be detectable, and actually *noticeable*, particularly in social situations. AI robots will know humans only as objects—active agents, but objects nonetheless. They will fail to empathize, to appreciate another social agent as the same as themselves. This is because models of self cannot be designed, but rather are dynamic and must be created from experience.

In contrast, developmental robots will have an internal model of "self" (their body space, their actions and effects, and their environment), and they will be able to match their experience of another similar agent or similar behavior with their own self-model and recognize a human as a separate agent: another "self." This allows much better interaction and social relationships because the robot can anticipate and appreciate why humans behave as they do. Learning will improve the model of the interactive partner, and exchanges will become more grounded in the context and understanding of both individual and joint intentions and goals.

After exploring these differences, I then take a somewhat wider look at the impacts that we might expect from this kind of technology. The inevitable spread of AI will affect all realms of human life. Such so-called disruptive technology offers both threats and benefits. I comment on some of the claims and alarms generated by the current research climate. This includes the threat to jobs, the role of trust, and the positive benefits of keeping humans involved in combined human/computer systems. It becomes clear that the threats from robotics and AI are part of the context of wider threats from general digital technology. In this part of the book, the observations spread out to apply to our total computing infrastructure.

The final chapters consider a number of projections for future progress and examine their feasibility. All sorts of nonsense is written in the media about robots, and it can be hard to know what to believe. Journalists love robotics stories because they allow huge scope for speculation,

A NOTE ON JARGON

I aim to avoid jargon, but I've already used the word *algorithm*! I will try to explain any overly technical terms that come up in the discussion. Actually, *algorithm* is a good example of a technical term that now has entered the vernacular. Once we talked about computer programs, then software, and now algorithms. At a colloquial level, all these can be taken as roughly equivalent, but to add a bit of discrimination, box 1.2 gives some details on the particulars of each term. Software usually consists of a collection of programs, and programs implement one or more algorithms; an algorithm is the essence of a computation. New terms have been coined for particular classes of application, as the last three illustrate.

I also try hard to avoid formulas and unexpanded acronyms—those dread, dead symbols that separate the technological laity from those "in the know."

For those who wish to follow up on some of the many fascinating topics, I include references to the relevant literature. I usually recommend either the classic papers or (for new concepts) introductory material, and other citations provide links to key technical sources.

2

COMMERCIAL ROBOTS

The *Encyclopaedia Galactica* defines a robot as a mechanical apparatus designed to do the work of a man. The marketing division of the Sirius Cybernetics Corporation defines a robot as: "Your Plastic Pal Who's Fun to Be With."
—Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, 1978

In 2012, something big happened. A team of artifical intelligence (AI) researchers made a major breakthrough: They discovered how computers could learn for themselves what features to use for automatically recognizing images. They smashed through a very big barrier that had hampered AI right from the beginning. AI systems could now learn tasks without handcrafting by their designers. This general area, using new learning methods and masses of data, has become known as *Deep Learning*.

Deep Learning generated much enthusiasm for AI and also raised expectations, especially in the media. We are currently in a boom period for AI and robotics; companies are heavily investing in these technologies, and research laboratories are creating a new wave of exciting applications and products.

This means that robots are becoming more useful, with new and better abilities, and becoming much more involved in everyday life. Unfortunately, the sheer number of innovations makes it harder to assess the current situation and evaluate the roles and value of robot systems.

We need a framework for identifying and understanding the fundamental issues among the wild diversity of forms and functions in this creative field. We need to concentrate on questions like: What do robots actually do? What are their limitations? What have they got in common? What classes of problem can they solve? Understanding these issues will give us a better sense of the way things are going and what to expect in the future.

The key feature of any robotic application is the nature of the *task* that the robot has to accomplish. In addition, there are *constraints* that have been either (1) avoided to simplify the task or (2) applied to control or restrict the working environment. For example, an agricultural robot is much easier to design if it needs only to spray weed killer over all the crops rather than to locate and pull out each individual weed (as some robots now do). This constraint makes the task easier (but not for organic farmers). Alternatively, an environmental constraint might be applied by structuring the field so that the weeds are easily identified (possibly by organizing or marking the crops). Constraints help to simplify things. The real world is full of chaos, noise, and uncertainty, and any constraints that a designer can exploit will significantly simplify things; using constraints effectively is a major talent of skilled engineers.

Let's start by looking at the design and behavior of a selection of commercial robots and assess their basic abilities. In each case, think about the tasks that they perform (what do they actually do?) and any constraints that they exploit. The answers will emerge by the end of chapter 3.

We should first define a robot as an autonomous physical device, of any shape or form, that can sense and perform actions in its working environment. Note that the requirement for autonomy rules out remote-controlled so-called robots such as drones.[1]

DOMESTIC ROBOTS AND SERVICE ROBOTS

Two tasks now performed by commercial robots are lawn mowing and floor cleaning. A Scandinavian company, Husqvarna, was the first to market a robot lawn mower, and it is now a major player with several models that vary according to the size of lawn;[2] as an example, see figure 2.1.

These small, battery-driven mobile devices can cut several thousand square meters of grass, entirely without human involvement. Batteries

2.1 A robot lawn mower, courtesy of Husqvarna.

provide less power than conventional mowers, and so they make many complete passes over the lawn, each giving a small trim, rather than one large cut. But this turns into an advantage because the smaller clippings don't need to be collected and are actually better for the lawn if left on the surface, similar to mulch. (It has been said that a really high-quality lawn is best cut with scissors, which is effectively what these robots do.) The mowers are quiet and run unattended; they determine their own cutting pattern, avoid objects, stay within a set region determined by a wire boundary marker, and return to their recharging station when they sense a low battery level. They can also use the Global Positioning System (GPS) to know where they are and can communicate with their owner via a smartphone. Robot lawn mowers are now standard products, offered in many garden suppliers' catalogs. They are made by a range of companies in France, Italy, Germany, the UK, and the US.[3]

Indoor floor-cleaning robots are also widely available and operate in a very similar manner. Instead of cutters, they use vacuum and brushes (for carpets) or wet mops (for hard floors). They cover areas bounded by walls, avoid falling down steps, and self-recharge. They are less expensive than mowers, and two of the main manufacturers, in Germany (Bosch) and the US (iRobot), are now well established in the global consumer market. This new market area has seen companies rush to bring many similar products to the market.[4] In the UK, six different brands were available in 2017, with the expensive (but very effective) Dyson 360 being the most highly rated.[5] Bigger robot cleaners are also available for large or public premises.[6]

These mowers and cleaners are both examples of *domestic robots*. Their specifications are very similar, involving the movement of a cutting or cleaning tool over a fairly well defined surface area.

They don't need legs (wheels work just fine), and the power and environmental requirements are not arduous. This means the task is well constrained (i.e., with very limited external interference) and not high in difficulty.

Another example based on the motorized mobile platform idea is the robot security guard. These robots wander around empty buildings or offices, like a human guard, and use various kinds of detectors, such as vision, ultrasonics, and infrared, to sense any humans or animals that might be present. They can then alert security staff through various networks and also may take video images of intruders. Again, a range of companies market this type of robot.[7]

A kind of role reversal of the security robot is the museum guide. These robots also work in public spaces, but they seek out people in order to give advice or assistance about the museum or venue. They have considerable entertainment value and tend to be developed ad hoc for particular places, and so they are not always commercial, but often rather experimental.

These indoor mobile robots are examples of *service robots*, which perform functions like the collection and delivery of office mail, medicines and supplies in hospitals, or robot servers in restaurants and couriers that deliver food. Most service robots have to interact with humans in some way and therefore need some social skills. This has usually been provided through fairly basic interfaces. Service robots have to deal with more complex environments than floor cleaners and so face somewhat more difficult tasks, while security robots face more complexity from dealing with unexpected human behavior.

FIELD ROBOTICS

Some of the most demanding robotic application areas are military and policing tasks, especially search and rescue and bomb disposal. The popular PackBot is a small, twin-tracked vehicle that can enter disaster areas, Daimler joined Audi and BMW in a deal to acquire Nokia's software-mapping company. This will be necessary to compete with Google's powerful mapping and geographic information systems.

Apart from known intense activity at the big car manufacturers, General Motors, Ford, Nissan, Renault, Toyota, Audi, BMW, Hyundai, Volvo, and Peugeot, there are smaller firms and start-ups that look very promising, such as Tesla Motors, nuTonomy, and Mobileye (recently bought by Intel). The emerging maturity of electric cars and hybrids, as well as the temptation for other high-tech firms like Apple to get involved, should accelerate development.

It is important to distinguish self-driving vehicles from completely driverless vehicles. There is a five-level taxonomy of autonomous vehicles (level 0 means no automation), ranging from level 1 (very minor assistance) to level 5 (full driving autonomy).[13] Most cars under development are at level 3 or 4, which requires a driver to be present, pay attention, and take control if needed. Vehicles with *no one on board* (level 5) are another matter entirely; the safety, legal, and ethical issues involved are much harder to solve.

Just assume for a moment that all the technical problems have been overcome; imagine sending your own vehicle to the supermarket to collect your order. Even though your personal robot car might be safe with regard to road and traffic behavior (although this is difficult enough), events might occur for which the system would be completely unprepared. Examples include taking verbal instructions from a farmer to avoid an accident, getting stuck in mud in the farmer's field, handling gridlock, dealing with vandalism, or just avoiding erratic drivers. Such events are almost certain to occur, but designers and manufacturers cannot consider *all* possible scenarios. This is known as a *long-tailed data problem*, wherein most of the data is clustered around a set of common cases, but there is a "long tail" of rare cases that are serious if they do happen.[14] The problem is that individually the long-tail cases are *extremely* rare on an individual basis, but in the aggregate, there are so many of them that there is a good chance of one occurring.

This might not matter in a constrained application area, such as a university campus, a factory complex, or a shopping mall, where simple procedures such as emergency shutdown and remote communication could be sufficient, but releasing unmanned robot vehicles "in the wild" (into the world at large) is sure to cause problems. In such situations, human drivers use their common sense and the background knowledge they have. AI research is working on this, but the problem now has nothing to do with driving and everything to do with general intelligence. Despite the manufacturers' confidence in ongoing research (http://www.daimler.com/innovation/en/) and their aim to commercialize fully autonomous cars so that "the car comes to the driver," it may take considerable time before these technical, ethical, and legal issues are sorted out to everyone's satisfaction.

MEDICAL ROBOTS

Medicine is another burgeoning application area for robotics, and although the equipment is expensive and tends to be very specialized, patients are now routinely offered robotic procedures for a range of treatments. For example, the da Vinci surgical robot provides minimally invasive surgery for prostate, bladder, and other internal organ operations (www.davincisurgery.com/). Other robotic devices assist surgeons in orthopedics and bone implants. Of course, these robots are not fully autonomous and surgeons are always in control, but they can provide more accuracy and better results than humans, especially where very high precision is required. For example, when fitting metal bone implants, the surface of the bone needs to be shaped to match exactly the implant that is to be glued onto it. The machining of the bone can be driven from computer data on the shape of the implant and reproduce the exact surface contours of the implant, thus producing a much better fit than human skill can achieve.

There are many other related applications, often known as *health-care robotics*, including hospital support robots and rehabilitation robots. But medical robotics is a complicated area. While some results are excellent, others involving soft tissue are not yet perfected, and the very high cost of medical equipment is a major concern.

SWARM ROBOTICS

Robot swarms are collections of many robots, often identical, that have to cooperate to perform tasks that are beyond the physical abilities of a single person or robot.

Robot soccer is a classic example of swarm robotics. The RoboCup competition (www.robocup.org/) has been held as an international tournament since 1997; it runs a series of leagues for different robots of different constructions and abilities, ranging from small LEGO robots, wheeled mobiles and larger designs, up to humanoids that, instead of just pushing the ball, actually kick it (!). This tournament series has been very successful in promoting and publicizing robotics and encouraged many schoolchildren and other amateurs to build their own robots.

In the early years, the robots had trouble finding the ball and not falling over, but progress has been impressive—although the stated target of holding a match between robots and humans by 2050 is an interesting open question! There are now many such contests involving robot designing and building, and they have given excellent publicity to the field.

Other forms of swarms include insectlike mini-robots, formation-flying aerial vehicles, and sea- and submersible-based swarms. As the numbers in a swarm increase, the demands on processing power in each agent decrease. Ants and termites build large structures and perform multiple functions, but they clearly have limited individual nervous systems.

An interesting case of an extreme robot swarm is the Kilobot swarm, developed at Harvard University (see figure 2.2). Kilobots are tiny circular robots, only 33 mm in diameter, which move in a kind of shuffle by vibrating their spindly wire legs. But they have an onboard processor, memory, and battery, and they can communicate with nearby Kilobots using infrared, as well as by sensing the distance between them. They also perceive ambient light levels and have their own controllable colored light output. The research aim is to explore how really large numbers of fairly simple agents can work together to achieve a task.

A large number of Kilobots can be programmed simultaneously with open-source software and modern development tools.[15] For example, in a foraging experiment. Images courtesy of Sheffield Robotics, University of Sheffield, 2018 (left) and Andreagiovanni Reina, Sheffield Robotics, 2019 (right).

Kilobots can perform self-assembly into a shape. The desired shape, essentially a silhouette, is given to all the robots, and they track round the edge of their neighbors and self-locate to form the shape. An identical program in each Kilobot contains a few simple procedures, like edge following. Other examples of this global-from-local behavior include foraging for food sources and returning to a designated nest, and following a leader in a snake formation.

A nice example of a practical swarm application is the automated robot warehouse. There are many kinds of automated warehouse: Some use conventional robot arms for loading pallets, some have high-speed gantries that access enormous arrays of storage space, while others use small trucks that navigate via various high-tech means around the floor of a warehouse. The latter approach often employs swarms of mobile robots to fetch and collect products from dispersed storage locations to make up a specific order. Perhaps the prime example is the Amazon robot warehouse, where small modular robots are able to move *under* a crate, jack it up off the floor, and carry it to the target destination. It is entertaining to see the orchestrated motion of dozens of robots chasing up and down the main "roads," entering side alleys, and lifting up and putting down their heavy loads. They whizz about the warehouse but, of course, never collide with each other or with people.

These robots were produced by Kiva Systems and are the brainchild of Raffaello D'Andrea, one of the company's cofounders.[16] Kiva Systems was set up in 2003, and some 80,000 of its robots are in use; Amazon found these little robots so valuable that they bought the company in 2012. Initially, it seemed that Amazon was going to enter the robot supplier market, but that didn't happen; Kiva has been absorbed into Amazon and renamed Amazon Robotics. There are some websites that contain interesting details on how the robots are used.[17]

ENTERTAINMENT ROBOTS

Another, very different, field is *entertainment robotics*. This is a large and diverse market. One of the first entries into this realm was a robot dog, AIBO (whose name stands for Artificial Intelligence roBOt), which was marketed by Sony between 1999 and 2005. The great thing about toys and entertainment is that these products don't have to address a particular task; they just have to be fun to play with! AIBO was provided with software tools that allowed the owner to program their 'pet' with its own behavior and personality. AIBO was very sophisticated, with a powerful processor, memory, and a good range of sensors: camera, tactile sensors, heat sensor, range finder, acceleration and vibration sensors, microphone, and speaker.[18] It also had some built-in programs for useful functions, such as righting itself if it was pushed over. AIBO was a platform for education as well as entertainment; it led the way to introducing children to engineering and science through fun workshops with robots. The RoboCup robot soccer competition, previously mentioned, held a special Four-Legged League, in which teams of AIBOs tracked an orange ball with their head cameras.

Surprisingly, after a gap of 11 years, Sony has released a new version of AIBO in 2018. It seems that Sony carefully reviewed the previous product and enhanced all the positive features. The new AIBO is more rounded, less robotic, with big eyes and cute appearance and behavior. It has more tactile sensing (it likes "a pat on the head, a scratch on the chin, or a gentle stroke down the back"), better vision and hearing, and can recognize people and commands. It has a curiosity drive that allows it to explore its environment, and it remembers familiar locations. This is very much and usually this rules out human ways of doing things, as they are often not the most efficient.

Therefore, looking at most commercial robots, we see highly focused, task-specific devices that perform very well at the particular task they were designed for but are quite unable to show the flexibility necessary for addressing other tasks. The key human characteristic of adaptability, which gives us such versatility, is notably absent. It is our ability to cope with *all* kinds of situations under all kinds of risk, uncertainty, and ignorance that is central to human achievement. You don't need human versatility to do repetitious, simple, or boring tasks, and hence machines are much better for such jobs.

Only the entertainment/companion robots show any tendency to behave *like humans*. This is because the need for dialogue is built into their task specification—they are designed to engage and play with their human owners. All the other robots have primitive user interfaces because that's all they need to communicate basic commands and signals. Robots like Cozmo and Kirobo Mini have to respond and react to people and are provided with programs that sequence appropriate actions from common behavioral patterns. However, they still lack the flexibility of humans. They do not really understand interactive exchanges and don't learn or adapt through social contact. They offer simple simulations of behavior, but they do not think or behave like humans.

OBSERVATIONS

- Application areas for robotics are extremely diverse, covering all possible activities. This leads to a wide range of very different devices.
- The physical form of a robot determines what it can do. Conversely, a given task may require a particular anatomy. Here's a crude example: Wheels are excellent for flat surfaces, while legs are better for rough ground and stairs.
- Robots are designed for particular tasks. An engineering task specification usually defines the function and structure of a device or machine.
- Constraints on the working environment make the problem simpler, and thus easier and less risky, and they allow for less costly solutions.

- The downside of good engineering is that the more efficient robots become for performing one particular task, the less flexible they are for doing anything else. Consequently, most robots do not perform at all outside their task specification. There are no general-purpose robots.
- Software is often embedded in robots, but remote links increasingly are being used to access very large data sources or complex AI systems.
- Social robots, mainly intended to engage with humans, are much less task-centered and will need to be more open-ended.

3

FROM RESEARCH BENCH TO MARKET

Which single word best characterizes science?
—Curiosity
Which single word best characterizes engineering?
—Creativity

Commercial robot offerings are not as remarkable as the new software coming out of artificial intelligence (AI) companies. Of course, robotics and AI are not the same thing, but robots increasingly adopt and use the new AI advances. So let's continue with our constraint assessment idea and look at some research robots.

Research robots are usually more sophisticated than commercial robots and are capable of demonstrating some very impressive behavior, but being experimental systems, they are often incomplete or dysfunctional in some nonvital way. It is useful to consider the research process model, which is rather different from that of much commercial development.

Where do robotic ideas originate, and how do they get tested? What are the drivers of innovation? Most commercial robotic products have their origins in upstream prototypes in robotics research laboratories. Such prototypes are usually not directly motivated by the marketplace, but rather are driven by curiosity and the search for new knowledge and understanding. Scientists are at their most productive when they explore new ideas

3.1 The Honda P3 humanoid robot, an early precursor of the ASIMO robot. This walking and stair-climbing robot stands 1.6 m (5.25 ft) tall and weighs 130 kg (286 pounds). Image courtesy of Honda Motor Corporation.

and try to gain a better understanding of basic scientific principles. Such principles are the bedrock of most of the engineering and technology that we all rely on. The downside is that they may sometimes find themselves in blind alleys, or their results may turn out to be negative or inconclusive, so their work does not always lead to new advances that can be immediately exploited. But this is a necessary condition of science. In fact, it is essential for the creation of an environment that can support breakthroughs and radical new advances.

But this creativity comes at a cost. For example, the Honda development program for its P3 humanoid robot (figure 3.1) took 10 years, and the estimated cost was over $10 million. The researchers started on legged locomotion with bipedal designs in 1986 and worked through a succession of improved experimental designs,[1] producing the P3 in 1997. This research-and-development (R&D) program paid off, as it provided the groundwork for the development of the impressive and world-famous ASIMO robot.[2] (But even for ASIMO, the payoff was not in profits from sales, but in advertising and public relations.)

3.2 A bin-picking task. Selected items are to be removed individually from a box of jumbled and very different products.

industrial and academic robotic communities and promote shared and open solutions to some of the big problems in unstructured automation. The Challenge will task entrants with building their own robot hardware and software that can attempt simplified versions of the general task of picking and stowing items on shelves. The Challenge combines object recognition, pose recognition, grasp planning, compliant manipulation, motion planning, task planning, task execution, and error detection and recovery. The Robots will be scored by how many items are picked and stowed in a fixed amount of time. Teams will share and disseminate their approach to improve future Challenge results and industrial implementations."[8]

Once a prototype solution has proved feasible and reliable, Amazon will implement it across the organization, thus gaining completely autonomous warehousing and massively reduced costs. This will repay many times over the few hundred thousand dollars the company has spent on the Challenge. This illustrates why technical challenges and competitions are popular with sponsors; they are a very effective and cheap way of stimulating innovation.

BIOROBOTICS

Much general inspiration for robotics research comes from the animal world. *Biorobotics* involves the selection of particular animal characteristics and behavior to reproduce through experimental implementations of the underlying mechanism. Examples that have been investigated include various birds and mammals, snakes, bats, and many forms of insects and reptiles, where locomotion, body dynamics, and specialized sensing regimes have all been studied. Robot fish include a robot tuna from the Massachusetts Institute of Technology (MIT) (Triantafyllou and Triantafyllou, 1995), an extinct coelacanth from Mitsubishi (Terada and Yamamoto, 2004), and three robotic fish from Essex University that swam in the London Aquarium (Hu, 2006). An interesting line of research at Vassar College is using robotics to understand the mechanics and dynamics of the spinal column of sharks, blue marlin, and tadpoles. The elasticity, stiffness, and other properties of the vertebral column in sharks are highly tuned by evolution for rapid and acrobatic swimming. But before the vertebrae had evolved, the earliest ancestors had backbones made from a solid flexible rod known as a *notochord*. It is not clear how these evolved and how the various biomechanics affected swimming performance, so the team at Vassar has been building robotic versions and measuring their performance while varying the mechanical parameters. Various arrangements of backbone give different energy storage properties and hence change dynamic behavior. This is another use of robotics: to model biological systems to understand the biomechanics of a current species and figure out how extinct species behaved.[9]

CARE AND ASSISTIVE ROBOTS

One of the big political worries for nearly all governments of the developed world is their aging population. Falling birth rates and increasing life spans mean that in the near future, there simply will not be enough younger people available to care for an enormously expanded population of older people, many with age-related impairments or disabilities. This is not an assumption, but a fact that we must face, demonstrated by demographic analyses of current populations.[10] Also, extended families are contracting in size and work patterns are changing for young people, exacerbating the growing shortage of caregivers available to provide support for aging populations. For these and other reasons, the care equation is going to become a crisis without careful planning. Over the last few decades, this concern has been a driver for supporting and funding robotics research in the social and care sectors. Governments and economists have come to believe that technology might provide a solution and have been motivated to fund large national research initiatives (e.g., in the US, Europe, and Japan).

There are two main care scenarios: helping people live independently in their own homes for as long as possible, and care homes, where people live together and centralized services can be provided. The former is the most desirable, from both the well-being and economics standpoints, but it is also the most difficult, as it involves dealing with every possible type of home and highly personalized needs and requirements. Indeed, this is perhaps the most taxing application area for robotics because so many different functions may be needed, and enormous flexibility will be required to adapt to various home scenarios. We have mentioned service robots before in this book, but they mainly transport items and deliver them inside hospitals or companies, whereas *assistive robotics* must support people by performing a wide range of tasks. Human caregivers typically visit their clients several times a day and help with getting dressed, preparing a meal, tidying the home, dealing with any problems, and then assisting with going to bed. It is difficult to imagine any current robot system being able to do some of these tasks, never mind all of them.

AFFECTIVE COMPUTING

Trying to build a completely humanoid robot that could reproduce the behavior of a human caregiver is *not* a sensible approach. So researchers are investigating various home tasks separately, including identifying and welcoming visitors; navigating around the rooms of a home; and finding and retrieving personal objects. Interactions with the user will be absolutely crucial: A good rapport and fluid communications are essential for acceptance, and this is obviously an area where social robots are needed.

Much work is underway in the specialized field of research investigating *human-robot interactions*. This covers the understanding of natural speech, dialogue and conversation, gesture and other forms of nonverbal communication, emotional behavior, and expression of meaning and intention. This is a very active research area and is producing valuable results thus far. For example, autistic children have shown responsive preferences for robot toys over human caregivers, and this stimulation and support from robots has enhanced their development (Wainer, Dautenhahn, Robins, and Amirabdollahian, 2010).

A particular topic is the study of empathy and human affective states. This is known as *affective computing*, or *emotion AI*, and concerns both the generation of emotional responses and the recognition of emotions in humans. Visual facial expression software can fairly accurately estimate human emotional moods by detecting changes in the facial muscles (McColl, Hong, Hatakeyama, Nejat, and Benhabib, 2016). Psychologists classify a range of emotions, but seven major categories are easily recognized by these systems: sadness, happiness, anger, fear, surprise, contempt, and disgust.[11] Emotions can also be detected from speech passages, or even textual material. Research robots have been built that can produce these expressions of common human emotions as a communication aid. Clearly, empathy is very important for social robots, but we will consider later (chapter 14) to what extent emotions are necessary.

HUMANOID ROBOTS

The most ambitious level of biorobotics is humanoid robots, where the aim is to implement models of the human form. Of course, these are the most overhyped kinds of robotics research; the image of a humanlike device raises expectations that it will exhibit humanlike behavior in every other way. There currently exists an enormous range of humanoid robots, ranging from build-it-yourself hobby kits, to affordable, multifunction commercial products, to expensive laboratory models used primarily for research. Several reviews are available,[12] and we will look at a few research milestones here.