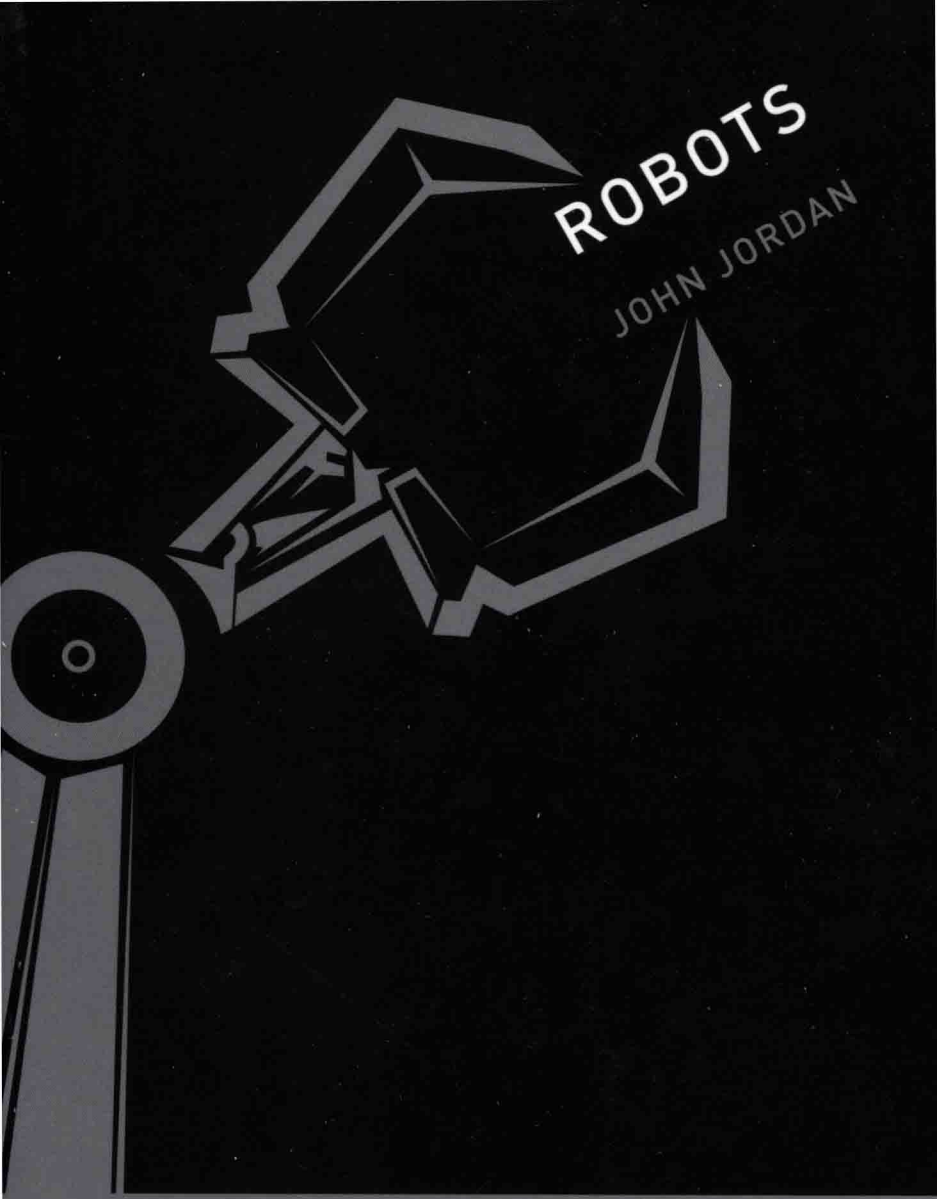

THE MIT PRESS ESSENTIAL KNOWLEDGE SERIES

ROBOTS

JOHN JORDAN

The MIT Press | Cambridge, Massachusetts | London, England

This book was set in Chaparral and DIN by Toppan Best-set Premedia Limited. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Jordan, John M. author.
Title: Robots / John Jordan.
Description: Cambridge, MA : MIT Press, [2016] | Series: The MIT Press essential knowledge series | Includes bibliographical references and index.
Identifiers: LCCN 2016008323 | ISBN 9780262529501 (pbk. : alk. paper)
Subjects: LCSH: Robotics—Popular works. | Robots—Social aspects—Popular works.
Classification: LCC TJ211.15 .J67 2016 | DDC 629.8/92—dc23 LC record available at https://lccn.loc.gov/2016008323

10 9 8 7 6 5 4 3 2 1

ROBOTS

The MIT Press Essential Knowledge Series

Auctions, Timothy P. Hubbard and Harry J. Paarsch
Understanding Beliefs, Nils J. Nilsson
Cloud Computing, Nayan Ruparelia
Computing: A Concise History, Paul E. Ceruzzi
The Conscious Mind, Zoltan L. Torey
Crowdsourcing, Daren C. Brabham
Free Will, Mark Balaguer
Information and the Modern Corporation, James W. Cortada
Intellectual Property Strategy, John Palfrey
The Internet of Things, Samuel Greengard
Memes in Digital Culture, Limor Shifman
Metadata, Jeffrey Pomerantz
MOOCs, Jonathan Haber
Neuroplasticity, Moheb Costandi
Open Access, Peter Suber
Paradox, Margaret Cuonzo
Robots, John Jordan
Self-Tracking, Gina Neff and Dawn Nafus
Waves, Frederic Raichlen

SERIES FOREWORD

The MIT Press Essential Knowledge series offers accessible, concise, beautifully produced pocket-size books on topics of current interest. Written by leading thinkers, the books in this series deliver expert overviews of subjects that range from the cultural and the historical to the scientific and the technical.

In today's era of instant information gratification, we have ready access to opinions, rationalizations, and superficial descriptions. Much harder to come by is the foundational knowledge that informs a principled understanding of the world. Essential Knowledge books fill that need. Synthesizing specialized subject matter for nonspecialists and engaging critical topics through fundamentals, each of these compact volumes offers readers a point of access to complex ideas.

Bruce Tidor
Professor of Biological Engineering and Computer Science
Massachusetts Institute of Technology

PREFACE

Attempting to write a book about robots represents a leap of faith. The field is already broad and expanding further, and it moves too fast for a multiyear project to be at all current. Why, then, did I proceed?

The field of robotics is, I believe, entering a crucial stage. Technologies are good enough for mass commercialization, use by governments, and even to become invisible in their ubiquity. Robots will soon affect millions of people's lives more directly and profoundly.

The technical choices made in robots' design not only embody value judgments and aspirations; they often have ethical implications. Every roboticist I have met is smart, humane, and well spoken. Even so, I do not want a small population of scientists and engineers, working in isolation, to make all the decisions that can affect life, death, health, work and livelihood, class status, personal privacy, gender identity, the future of war, the shape of urban landscapes, and many other domains: they need help, and other perspectives.

This book seeks to widen the circle of individuals who have a say in what robots can and should do, look like, include, and leave out. I hope roboticists read the book, but the intended audience is the rest of us. Design choices made now may well be with us for decades, so now is the

time to ask, and to assert, what "good" robots might look like. Because of the breadth and dynamism of the field, my narrative's emphasis falls less on completeness and the latest developments, and more on the enduring issues: what are the persistent capabilities, contests, and trade-offs that will characterize robots and robotics?

Why do these issues matter? Many powerful scenarios involve humans and robots working in partnership rather than robots *replacing* humans, whether on a battlefield, in a hospital, on an assembly line, or in the rehabilitation, prosthetics, or aging processes. Rather than focusing on either-or debates over what constitutes a robot, we will be better served by seeing compu-mechanical augmentation of human traits on a continuum. By necessity, this implies both that robots and human beings will be living and working closely, changing the nature of the human condition in important ways, and that, rather than becoming only slaves or potentially overlords, robots will become humanity's partners. These impending changes make improving our theories, norms, and aspirations an urgent matter. This book, this leap of faith, is a small step in that direction.

ACKNOWLEDGMENTS

This slim volume, five years in the making, was only possible with the help of many people, not all of whom can be named here.

Katherine A. Almeida, Kate Hensley, and especially Marie Lufkin Lee at the MIT Press have been consummate professionals, helping me, by turns, with advice, encouragement, constructive challenges, and superb execution. I count myself fortunate to have had the backing of such a capable team.

When Bob Bauer showed me the personal robot PR2 at Willow Garage back in 2011, that was the precise moment—on a par with my first glimpse of the National Center for Supercomputing Applications (NCSA) Mosaic web browser—that I saw everything differently. Bob later introduced me to my key research interviewees, including Steve Cousins, Scott Hassan, James Kuffner, and Leila Takayama; without him, this book would never have happened, so my thanks are profound.

In addition to many anonymous readers from MIT Press who each improved this book, I can thank four readers by name. Steve Sawyer offered a generous, incisive critique of an earlier draft and dazzled me with the breadth of his suggestions. Kate Hoffman deepened my understanding of anime and science fiction, drawing on her vast trove

of reading recollections and judgments. My longtime partner in tech analysis, John Parkinson, read multiple drafts and unfailingly helped me, whether by cracking macro framing problems or correcting imprecision in the details.

Finally, my onetime coauthor, David Hall, offered encouragement, critical reaction, and expert perspective every time I asked him, even though he had dozens of other irons in the fire. Death took him, suddenly and too soon, a month before this book went into production. I can only hope he would have been proud of the final result.

CONTENTS

INTRODUCTION

Television and cinema have helped propagate powerful American stories of technology. The *Star Wars* franchise is in many ways an updated Western, using space as a frontier (though maybe not the final one) and light sabers as rifles. The impact of such icons—and possibly archetypes—as Robocop, *Blade Runner*'s replicants, the brilliant, mannerly C-3PO, and Disney's Wall-E is wide and deep. Ask about "throwbots," Atlas, Motoman, Kiva, Beam, or other real-life robots that are reinventing warfare, the industrial workplace, and human-robot collaboration, and most people have little sense of what real robots do or look like. But everyone knows the Terminator, complete with Austrian accent.

In 2004, Chris Van Allsburg's beloved children's book *The Polar Express* was made into a movie. Oscar-winning

talent (including Tom Hanks) was wired up for digital motion capture, but the resulting characters came out, to use critics' words, "creepy," "eerie," and "dead-eyed"; the movie was "a zombie train." For years, digital animators pushed for more polygons, more shades of color, more pixels—in short, more computing. But instead of delighting audiences with verisimilitude, they entered the "uncanny valley," a paradox in which added artificial likeness weirds people out. After 2005, the Japanese female android Repliee generated the same reaction with its almost-but-not-quite humanlike features.

In contrast to Hollywood, Boston Dynamics (between 2013 and 2016, a Google company) builds robots intended for use by the U.S. armed services. Nonclassified YouTube videos of a robotic cheetah, humanoid, and pack animal have received tens of millions of views, giving many people their first look at state-of-the-art robotic science. Rather than the moderately large viewership, most striking to me was my students' reaction when a human shoved/kicked the BigDog to show its stability: they gasped, as if someone had struck a pet on camera.

Robots are becoming more numerous, more capable, and more diverse. Over the long term, their economic, civic, and destructive impact will likely be on a par with that of the automobile. In such a massive transition, people will care what happens and call for rules, norms, and paths of recourse. Citizens have a vested interest in

work, wages, and workplace safety; in aging with dignity; in major changes in global warfare; in privacy; and in other things that robotics has the potential to change. Many factors, however, combine to make it hard to advance the discussions about what we want from today's and tomorrow's robots.

Barriers to Informed Dialogue

When an innovation emerges, the history of its naming shows how it goes from foreign entity to novelty to invisible ubiquity. A little more than 100 years ago, automobiles were called "horseless carriages," defined by what they were not rather than what they were. More recently, the U.S. military referred to drones as "UAVs," unmanned aerial vehicles, continuing the trend of definition by negation.

The word "robot" originated in the 1920s and was at first a type of slave; robots are often characterized by their capabilities in performing dull, dirty, or dangerous tasks, sparing humans the need to perform them. The science and engineering of the field continue to evolve rapidly—look no further than Google's self-driving car or the humanoid robots it acquired in the Schaft and Boston Dynamics deals. Given such rapid change, computer scientists cannot come to anything resembling consensus on

what constitutes a robot. Some argue that a given device qualifies if it can (1) sense its surroundings; (2) perform logical reasoning with various inputs; and (3) act upon the physical environment. Others insist a robot must move in physical space (disqualifying the Nest thermostat), while still others say that true robots are autonomous (excluding factory assembly tools). *Reason 1 why robots are hard to talk about: the definitions are unsettled, even among those most expert in the field.*

Bernard Roth, a longtime professor of mechanical engineering who was associated with the Stanford Artificial Intelligence Laboratory (SAIL) from its inception, offers a more nuanced definition that draws on his long history in the field. Roth begins by doubting that "a definition [of what is or is not a robot] will ever be universally agreed upon." Instead, he argues for a much more relative and conditional approach: "My view is that the notion of a robot has to do with which activities are, at a given time, associated with people and which are associated with machines." When relative capabilities evolve, so do conceptions: "If a machine suddenly becomes able to do what we normally associate with people, the machine can be upgraded in classification and classified as a robot. After a while, people get used to the activity being done by machines, and the devices get downgraded from 'robot' to 'machine.'"[1] *Reason 2: definitions evolve, unevenly and jerkily, over time as social context and technical capabilities change.*