Auton Agent Multi-Agent Syst
DOI 10.1007/s10458-008-9041-1

Book Review: Maja J. Matarić: The Robotics Primer

MIT Press, Cambridge, 2007, 300pp, $30.00, ISBN 0-262-63354-X

Jason M. O'Kane

After decades of development, truly autonomous robots are finally breaking through into everyday life. Many observers (including, for example, Bill Gates [3]) expect this progress to spur societal changes on a scale comparable to the information revolution of the past 20 years. To this context, Maja J. Matarić brings *The Robotics Primer*, a broad, introductory-level robotics textbook aimed at a diverse audience, including "K-12 and university educators, as well as robotics enthusiasts who want to go beyond the popular press and get under the surface of the topic," spanning "the ages of early adolescence to retirement, and education levels of elementary school to Ph.D." The stated goal of the text is to "teach something real and true about robotics and its potential, while staying interesting and engaging," and in this regard it is a very successful book. The end result is a detailed guided tour of a large, interdisciplinary field.

Although many fundamental robotics concepts are most precisely described with mathematical language, *The Robotics Primer* relies almost exclusively on conceptual, rather than overtly mathematical descriptions. The result is that many topics (inverse kinematics, underactuated systems, machine learning, and many others) are made accessible, at a basic level, to students much earlier in their academic careers, and to experts in other fields more easily than in existing books and survey papers. One hopes that readers are inspired to learn the details eventually, rather than being discouraged by the occasional warning that they should "be glad for not having the math included here."

Equally important, the conversational style of the text makes it generally very readable. A telling example is that in the first chapter, the book describes itself as "A very cool way to learn about a very cool topic." However, in a very few cases, this kind of informal language may lead to confusion. For example, the frequent reference to the robot's "brain" and to what the robot "thinks" may hinder non-expert readers from appreciating the difference between strong AI efforts to create machines with human-level intelligence and the more modest objective of simply building useful autonomous devices. Other instances suffer from uncomfortable leaps in vocabulary. One wonders, for example, how many novice readers will

J. M. O'Kane (✉)
Computer Science and Engineering, University of South Carolina, Columbia, SC 29208, USA
e-mail: jokane@cse.sc.edu

understand the description of analog signals as being "continuous in time and amplitude." The chapters are generally short, focused, and packed with a dense network of internal references. Overall, the text succeeds in its goal of being accessible to younger and less technical readers.

It should also be noted that *The Robotics Primer* is much more than simply a programming guide or "cookbook." Matarić takes great care to discuss the options and engineering tradeoffs that so strongly govern the art of robot design. For example, one section examines seven distinct methods for sensing the distance to the nearest obstacle, weighing the merits and demerits of each. Although specific real hardware is described (including Polaroid sonars, Sick laser range scanners, and others), the focus is always on the concepts underlying the technology and its use in robotics. This focus makes the book a lasting contribution that will not need to be replaced after the next inevitable wave of technological advances. Indeed, the book is meant as a rudimentary introduction to the principles of robotics, emphasizing links to more advanced resources, including the popular undergraduate textbooks of Murphy [5] and Arkin [1], more recent books like those of LaValle [4] and Choset et al. [2], and even a few Ph.D. theses [6,7]. Naturally, most of this suggested material will be of more immediate value to some members of the book's audience than to others.

The text is divided into 22 chapters that can be informally grouped into three parts. The first part spans nine chapters and provides introductory material, along with an inventory of the essential physical components that comprise a complete robot. Chapter 1 unpacks Matarić's definition of a robot as "an autonomous system which exists in the physical world, can sense its environment, and can act on it to achieve some goals." The requirement of full autonomy is forward-looking and important, but also discounts somewhat the exciting possibility that robots and humans working together might be more effective than either one alone. Chapter 2 is a short history of robotics, emphasizing the field's interdisciplinary roots, and particularly the contrast between deliberative and reactive designs that persists even today.

From this basis, the text moves to a discussion of robot components, first at a very high level (Chapter 3), then in greater detail. The material on actuators (Chapter 4) covers a surprising breadth of options, well beyond simple motors and pneumatics, and also nicely illustrates the basic issues that arise in controlling underactuated systems. Locomotion (Chapter 5) and manipulation (Chapter 6), the two fundamental goals of actuation systems, are covered separately. The discussion of sensing (Chapter 7) emphasizes the tradeoffs between the information content provided by sensors and the processing needed to utilize them well. Matarić classifies sensors as either simple (useful with very little processing, Chapter 8) or complex (requiring nontrivial processing, Chapter 9). Even some simple ideas from machine vision appear, a welcome surprise to see at such an accessible level. Overall, the view that robot design is an exercise in tradeoffs was much appreciated and well-supported.

The second major section presents architectures for controlling robots, focusing on four major choices: deliberative, reactive, hybrid, and behavior-based architectures. First, however, Chapter 10 presents some basic concepts from control theory, devoting the greatest attention to the intuition behind PID control. This is the only significant time Matarić uses mathematical notation. It is not made clear, however, how these notions from control theory are related to the reactive control architecture introduced later. Definitions and motivations for using control architectures to organize robot programs appear in Chapter 11. The statement that all "effective robot programs" fit into one of four categories "even if the programmer does not realize it" seems suspicious on the surface, but is reasonable if the categories are interpreted broadly enough. Chapter 12, which discusses representation issues, gives an example of several different representations (metric, action-based, landmark-based, and topological) for an environment map. Unfortunately, Matarić misses the chance to tie representation more

closely to the sensing capabilities of the robot: the sensing available to the robot determines the models that are sensible to attempt to build, and conversely, the models desired can influence the robot's sensing design.

Next, Matarić presents each of the previously mentioned control architectures, starting with the deliberative control architecture (Chapter 13). The tone here seems somewhat dismissive, suggesting that deliberation has little long-term place in robotics. Much of this criticism seems motivated by the well-known limitations of early, AI-inspired systems like Shakey, whose usefulness was severely hampered by computational limitations. Many of the author's criticisms of planning approaches are legitimate, but the overall approach is dismissed too quickly. In particular, it would have been interesting to see some consideration of newer, more flexible notions of planning that anticipate changes, replan in the face of exceptions, and reason explicitly about missing knowledge. The issue of representation choice is important here as well, since minimal, streamlined representations lend themselves much better to efficient planning. In contrast, reactive control (Chapter 14) eschews representation in favor of direct mappings from sensor readings to actions, but can require prohibitively large rulesets for many problems.

To balance the strengths and weaknesses of these two architectures, Matarić discusses two intermediate possibilities. Hybrid control (Chapter 15) uses both deliberative and reactive layers, with their interactions governed by a middle layer. Behavior-based control (Chapters 16–17), in contrast, describes the robot's control scheme as a collection of interacting behaviors, avoiding explicit, centralized representations and deliberation. Although the author claims that the choice between hybrid and behavior-based architectures is a matter of preference, the motivation for behavior-based control, which appears counterintuitive in some of its details, seems somewhat unconvincing.

The final major section is a collection of advanced topics that reach beyond the basic model on which the remainder of the book focuses. Chapter 18 spends significant time developing a definition of emergent behavior. In the end, this effort seems misplaced—far more relevant than whether or not a given behavior is "emergent" is whether the system achieves its goals—but modern roboticists certainly will encounter the term, and perhaps they should be introduced to the concept, regardless of their views on its importance. Chapter 19 collects algorithms for coverage, localization, and mapping under the general heading of "navigation." These topics are important and, in some sense, are at the heart of algorithmic robotics, but the chapter seems misplaced in the overall bottom-up organization of the book, particularly since the methods described here are essentially deliberative and representation-based.

Of greatest interest in this final section is the discussion of multi-robot systems (Chapter 20). Matarić gives a broad taxonomy of multi-robot systems: heterogeneous versus homogeneous, amounts and types of communication, and levels of coupling between robots. Since the ultimate goal of many efforts in multiagent systems is to build autonomous physical systems, the content of this chapter may provide a starting point for examining how issues unique to robotics should influence the design of multiagent systems, and likewise how the overall design of a multiagent system should influence the construction of the individual robots.

The final chapters discuss learning issues in the context of robotics (Chapter 21), followed by a summary of the author's vision of the future of robotics (Chapter 22).

Finally, several other, more global comments are in order. First, at several points in the first half of the book, the presentation seems piecemeal, making it difficult to see the "big picture" of a completed robot. Fortunately, the book is accompanied by a tightly-integrated free online workbook (http://roboticsprimer.sourceforge.net) that has the potential to fill in

the gaps, especially in classroom settings. For approximately half of the chapters, it provides hands-on exercises based on the iRobot Create platform. The workbook is advertised as a community effort, which, at the time of this writing, is still developing. Many of the exercises are tersely written and only some have solutions provided. It seems reasonable to expect this companion material to continue to improve over time.

Second, Matarić's viewpoint places strong importance on biological inspiration for robot systems. When, in the discussion of behavior-based control, the author writes that "all approaches to control ... were inspired by biology," one is tempted to wonder whether this statement illustrates a desire to see links to biology more strongly than they actually exist. Admittedly, living creatures are the only known examples of systems that meet Matarić's definition of a robot in its strongest sense, but it is a matter of debate whether designing robots to mimic biological systems is the most effective course, especially in the short term. The book would benefit from a broader view that includes this debate.

Third, the notion of a "state"—a crucial concept for reasoning about how robots interact with their environments—is unclear at several points. In one place, the author reminds us that "sensors do not provide state," but elsewhere (for example, in Fig. 3.2 and in Sect. 21.1) we see examples in which the robot's state space is defined as the set of all possible immediate sensor readings. As a result, it may be difficult for readers to understand the difference between the state space and the sensor space. The concept of partial observability, which would nicely reconcile these two views, is mentioned very early, but never reappears.

In spite of these relatively minor flaws, few books deal as comprehensively with the interactions of mechanical, electronic, and computing issues that lie at the heart of robotics. None do so in such an accessible way. As such, *The Robotics Primer* will serve for many years as an introduction to the field for both young people and for nonexperts.

References

1. Arkin, R. C. (1998). *Behavior-based robotics*. Boston, MA: MIT Press.
2. Choset, H., Lynch, K. M., Hutchinson, S., Kantor, G., Burgard, W., Kavraki, L. E., & Thrun, S. (2005). *Principles of robot motion: Theory, algorithms, and implementations*. Boston, MA: MIT Press.
3. Gates, B. (2007). A robot in every home. *Scientific American Magazine*, January 2007.
4. LaValle, S. M. (2006). *Planning algorithms*. Cambridge, U.K: Cambridge University Press. Available at http://planning.cs.uiuc.edu/.
5. Murphy, R. (2000). *Introduction to AI robotics*. Boston, MA: MIT Press.
6. Pirjanian, P. (1998). *Multiple objective action selection and behavior fusion using voting*. Ph.D. thesis, Aalborg University, Aalborg, Denmark.
7. Rosenblatt, J. (1995). *DAMN: A distributed architecture for mobile navigation*. Ph.D. thesis, Carnegie Mellon University Robotics Institute, Pittsburgh, PA.