11 Embodiment in Cognitive Science and Robotics

Tom Ziemke

11.1 Introduction

Writing a book chapter on the notion of *embodiment* in the cognitive sciences, or cognitive robotics more specifically, is not an easy task these days. Many researchers nowadays share the belief that, as M. Wilson (2002, 625) formulated it, "Cognitive processes are deeply rooted in the body's interactions with the world," and we can take that as a useful first approximation of the fundamental claim of *embodied cognition* research. That statement, however, means surprisingly many and surprisingly different things to different people.

Hence, explaining what embodiment is, in a single chapter, is difficult for several reasons: First, embodiment has been discussed in the cognitive sciences for several decades now. Early examples include Lakoff and Johnson's (1980) work on the role of bodily metaphors in human cognition and language, as well as Maturana and Varela's (1980) work on the biology of cognition. Moreover, many of these discussions have roots preceding cognitive science as a discipline by several more decades, such as the work of William James in the late nineteenth century. So there simply is a long history to cover. Second, embodied cognition has become a popular and more-or-less mainstream position in the last twenty years (e.g., Clark 1999; Damasio 1999; Gallagher 2005; Ziemke et al. 2006; Pfeifer and Bongard 2007; Chemero 2009; Shapiro 2010; Black 2014; Lindblom 2015). Some would go so far as to claim that "embodied cognition is sweeping the planet," as it says in one of the endorsements on the back cover of Shapiro's (2010) textbook on the topic. However, research on embodied cognition is not so well established and mainstream yet that research has converged sufficiently to establish clear boundaries and shared definitions of what is or is not embodied cognition (e.g., Wilson and Golonka 2013; Ziemke and Thill 2014; Ziemke 2016). Hence, there are many different—and in some cases also conflicting—perspectives to address. Last, but not least, the issue of embodiment is somewhat uniquely placed at the intersection of engineering, science, and philosophy, which means that embodiment simply has different significance to different, but overlapping, research communities.

The latter point should be relatively clear in the context of a book on cognitive robotics (see also section 1.2): On the one hand, there is the *engineering perspective* (with an emphasis on the "robotics" in "cognitive robotics") on how to equip robots with the required sensorimotor, cognitive, and communicative capacities for particular tasks. If you think of a typical example, such as a service robot helping elderly people at home, it is clear that the robot's embodiment—in the sense of what the robot looks like, what sensors, actuators, and interactive modalities it has, and so on—plays a crucial role in determining what it can do and how people can interact with it. After all, robot lawnmowers and vacuum cleaners, for example, might be well suited for their specific purposes, but they are not exactly easy to talk to. On the other hand, there is the (cognitive-) *scientific perspective* (with an emphasis on the "cognitive" in "cognitive robotics"), according to which embodied (i.e., robotic) models that share some bodily and sensorimotor features with the organism they are supposed to model might be preferable to purely computational or mathematical models. If, for example, you are working on modeling how human language use is grounded in sensorimotor interaction, then it might make sense to use humanoid robot models that are, at least to some degree, similar in terms of bodily features and sensorimotor capacities to the people and processes you are trying to model. On the third hand (to use an intentionally confusing bodily metaphor), there is of course the *philosophical perspective*, according to which theories of embodied cognition and cognitive-robotics models offer novel approaches to age-old questions concerning the so-called mind-body problem.

To cover a broad range of perspectives on embodiment, the remainder of this chapter is structured as follows: section 11.2 asks some basic questions—such as what is a body, what do we mean by embodiment, and what do we mean by embodied cognition—and provides some preliminary answers in the form of basic distinctions that might be useful. Section 11.3 then addresses fundamental conceptions of embodied cognition in cognitive science and the philosophy of mind. Section 11.4 narrows the perspective to notions of embodiment in artificial intelligence (AI) research, where naturally some of the central questions are what would constitute an artificial body or embodiment capable of supporting artificial embodied intelligence, and how we should go about building such systems. Section 11.5 then addresses the role of embodiment in cognitive robotics more specifically and connects back to the above discussion of different perspectives on embodiment. Section 11.6, finally, provides a brief summary and some conclusions.

It should be noted that throughout this chapter, for the reasons mentioned above, we prioritize breadth—that is, provide a broad spectrum of perspectives on embodiment and its role in human (and robotic) cognition and refer the interested reader to the original literature for more in-depth discussions.

11.2 Notions of Embodiment and Embodied Cognition

What actually is a body? Or, more specifically, what constitutes the kind of body—or embodiment—that might be a necessary requirement for embodied cognition? Somewhat surprisingly, maybe, many discussions of embodied cognition actually pay relatively little attention to the nature and the role of the body involved. This might be natural in psychology or linguistics, which mainly deal with phenomena such as human cognition or language, where it is more or less obvious that discussions of embodiment are about the role that *human bodies* play in such phenomena. In AI and robotics, however, things are less obvious. This raises questions, such as what kind of embodiment might be required for an artificial system that could deal with, for example, human language. In the realm of science fiction, you might have noticed that the *Star Wars* android C-3PO claims to be fluent in six million forms of communication and thus is presumably capable of dealing with many different species, although his embodiment is rather humanlike. In today's real world, on the other hand, many of us regularly encounter systems that appear rather disembodied, such as Google Translate, or devices with somewhat minimal and distinctly nonhuman physical "embodiments," such as Amazon Echo or Google Home, which nevertheless all seem to be able to deal with human language to some degree. Let us therefore have a quick look at some of the notions of what kind of body or embodiment is required for embodied cognition (following Ziemke 2001, 2003b).

Embodiment as Structural Coupling

Probably the broadest notion of embodiment is that systems are embodied if they are structurally coupled to their environment. The concept of structural coupling originally comes from Maturana and Varela's (1980, 1987) work on the biology of cognition, which will be discussed in further detail below. Quick and colleagues (1999, 340) tried to formalize this as follows in their minimal definition of embodiment: "A system X is embodied in an environment E if perturbatory channels exist between the two. That means, X is embodied in E if for every time t at which both X and E exist, some subset of E's possible states with respect to X have the capacity to perturb X's state, and some subset of X's possible states with respect to E have the capacity to perturb E's state." This definition of embodiment in itself does not make a distinction between cognitive and noncognitive systems, which is illustrated by Quick et al.'s (1999) example of a granite outcrop on the Antarctic tundra that is perturbed by the wind and in turn perturbs the flow of air. This would seem to include practically all physical objects, but it might be worth noting that it has been argued that structural coupling does not necessarily require a physical body. Franklin (1997, 500), for example, explicitly stated: "Software systems with no body in the usual physical sense can be intelligent. But they must be embodied in the situated sense of being autonomous agents structurally coupled with their environment."

Historical Embodiment

Several researchers have emphasized that cognitive systems are not only structurally coupled to their environment in the present. Their embodiment is in fact a result or reflection of a history of agent-environment interaction. According to Varela et al. (1991, 149), for example, "Knowledge depends on being in a world that is inseparable from our bodies, our language, and our social history—in short, from our embodiment." Ziemke (1999, 187) pointed out: "Natural embodiment is more than being-physical . . . it reflects/embodies the history of structural coupling and mutual specification between agent and environment in the course of which the body has been constructed." In a similar vein, Riegler (2002, 347) included an agent's adaptation to its environment over time in his definition of embodiment: "A system is embodied if it has gained competence within the environment in which it has developed."

Physical/Sensorimotor Embodiment

Many researchers in embodied (robotic) AI—to distinguish their approach from traditional AI—hold that, as Pfeifer and Scheier (1999, 649) formulated it, "intelligence cannot merely exist in the form of an abstract algorithm but requires a physical instantiation, a body." This would seem to rule out software agents but could possibly still be considered to include the granite outcrop mentioned above. However, most researchers in embodied AI and robotics naturally adopt a more restrictive version of the notion of physical embodiment—which might be labeled *sensorimotor embodiment*—that is, the view that embodied systems need to be connected to their environment not just through physical forces but more specifically through their own sensors and actuators. This is also the essence of Brooks's (1990) *physical grounding hypothesis*, according to which building an intelligent system requires having its representations grounded in the physical world, which in turn requires connecting it to the world via sensors and actuators.

At this point it might be worth pointing out that, although historical embodiment and physical/sensorimotor embodiment can be considered special cases of structural coupling, neither of these notions necessarily includes or excludes the other. Riegler (2002), for example, stated that his historical notion definition of embodiment "does not exclude domains other than the physical domain" and in particular that "computer programs may also become embodied" if they are the result of self-organization rather than conventional human design and programming.

Organismoid Embodiment

Another, yet more restrictive, notion of physical and sensorimotor embodiment would be that at least certain types of organism-like cognition might be limited to organism-like bodies—that is, physical bodies with particular structural features or sensorimotor capacities. A simple early example of this was a robot used by Lund et al. (1998) equipped with an auditory circuit and two microphones the same distance from each other as the two "ears" of the cricket whose phonotaxis it was supposed to model. The placement of the sensors/ears, in both robot and cricket, reduced the amount of internal signal processing required to respond selectively to certain sound frequencies. Note that in this case the bodies of the cricket and the wheeled robot were of course very different except for one crucial detail, namely the distance between the "ears." More recent and complex examples of organismoid embodiment can be found in humanoid cognitive robotics, such as the work of Cangelosi and colleagues, in which humanlike embodiment is taken to be a key ingredient for robotic models of human language learning and use (e.g., Morse et al. 2015; cf. section 11.5).

Organismic Embodiment

The most restrictive notion of embodiment discussed in this section is that embodied cognition emerges from the interaction between organisms—that is, living bodies, and their environments. Maturana and Varela's (1980, 1987) work on the biology of cognition, for example, suggests, in a nutshell, that cognition is what living systems do in interaction with their environment. In a similar vein, from a neuroscientific perspective Damasio (1998) criticized "the prevalent absence of a notion of organism in the sciences of mind and brain" as a problem, which he elaborated as follows: "It is not just that the mind remained linked to the brain in a rather equivocal relationship, but that the brain remained consistently separated from the body and thus not part of the deeply interwoven mesh of body and brain that defines a complex living organism" (Damasio 1998, 84). Somatic theories of emotion constitute "a multi-tiered affectively embodied view of mind" (Panksepp 2005, 63), according to which emotion, cognition, and consciousness arise from multiple, nested levels of homeostatic (self-) regulation of bodily activity. This is, at least at this point in time, a clear difference between living systems and robotic bodies, which typically have no *own* needs or viability constraints (e.g., Bickhard 2009; Ziemke 2016) and therefore need to be equipped with artificial motivational systems (cf. chapter 13).

In addition to the above different notions of embodiment, let us also take a quick look at different views of embodied cognition. M. Wilson (2002) distinguished six such views from a psychological perspective, of which, however, only the last explicitly addresses the role of body:

"Cognition is situated" This claim is widely held in the literature on embodied cognition. M. Wilson (2002) herself distinguished between situated cognition, taking place "in the context of task-relevant inputs and outputs" and "off-line cognition."

"Cognition is time pressured" Cognition is constrained by the requirements of real-time interaction with the environment, such as the representational bottleneck (e.g., Brooks 1991; Clark 1997; Pfeifer and Scheier 1999).

"We off-load cognitive work onto the environment" Brooks (1990) formulated a similar claim stating that "the world is its own best model." Another well-known example is Kirsh and Maglio's (1994) study of the *Tetris* computer game players' epistemic actions—that is, decision-preparing movements carried out in the world, rather than in the head.

"The environment is part of the cognitive system" The classical example is Hutchins's (1995) work on distributed cognition, considering, for example, the instruments in a cockpit as parts of the cognitive system.

"Cognition is for action" A claim made, for example, by Franklin (1995), who argued that minds are the control structures of autonomous agents.

"Off-line cognition is body-based" According to M. Wilson (2002, 625), at the time this claim had received the least attention in the cognitive science literature, although "it may in fact be the best documented and most powerful of the six claims." An early example is the aforementioned work of Lakoff and Johnson (1980, 1999) who argued that abstract concepts are based on metaphors grounded in bodily activity and experience (such as the English expression "to grasp," which refers to both manual grasping of physical objects and the more abstract grasping of, for example, ideas or concepts). More recently, the underlying mechanisms have been elaborated in terms of *embodied simulation* accounts of conceptualization and cognition (e.g., Gallese and Lakoff 2005; Gallese 2005).

It might be worth noting here that, in one way or another, all of the above six views deal with the sensorimotor interaction between an agent's body and its environment, but none addresses the question of whether the body involved necessarily needs to be physical, biological, and so on. Again, from the perspective of psychology or linguistics—with a focus on human cognition and language—this might be understandable, and the physical or biological nature of the body involved might be considered a nonissue. However, Johnson's (2007, x) account of the development of research on the embodiment of language, which also initially focused on the sensorimotor aspects, indicates that the underlying biological mechanisms were initially somewhat overlooked: "In retrospect I now see that the structural aspects of our bodily interactions with our environments upon which I was focusing were themselves dependent on even more submerged dimensions of bodily understanding. It was an important step to probe below concepts, propositions, and sentences into the sensorimotor processes by which we understand our world, but what is now needed is a far deeper exploration into the qualities, feelings, emotions, and bodily processes that make meaning possible."

11.3 Embodiment in Cognitive Science

The introduction referred to M. Wilson's (2002) general statement that "cognitive processes are deeply rooted in the body's interactions with the world" as a useful first approximation of the fundamental claim of embodied cognition research. Different general notions of embodiment and embodied cognition have already been addressed in the previous section. In this section, let us quickly look at the general theoretical conceptions of embodied cognition in cognitive science and in particular philosophy of mind. The following somewhat simplified diagram from Chemero (2009) provides one useful perspective on the current embodied cognition research landscape (figure 11.1). As Chemero pointed out, there are at least two rather different general theoretical frameworks that are both referred to as "embodied cognitive science." One of these, which Chemero referred to as "radical embodied cognitive science," is grounded in the antirepresentationalist and anticomputationalist traditions of eliminativism, American naturalism (such as the work of James and Dewey), and Gibsonian ecological psychology. The more mainstream version of embodied cognitive science, on the other hand, is derived from the traditional representationalist and computationalist theoretical frameworks that have long dominated cognitive science, and therefore is still more or less compatible with these. In a cognitive robotics context, the latter can be exemplified with the popular notion of symbol/representation grounding (Harnad 1990; Ziemke 1999), whereas the former is closer to the antirepresentationalism advocated by embodied AI researchers such as Brooks (1991) and Beer (1995). As Chemero pointed out, although—or maybe just because—the mainstream version of embodied cognitive science can be considered a "watered-down" version of its more radical counterpart, it currently receives significantly more attention in the cognitive sciences.

Chemero's (2009) formulation of radical embodied cognition can be summarized in the following claims:

1. Representational and computational views of embodied cognition are wrong.
2. Embodied cognition should be explained using a particular set of tools *T*, including dynamical systems theory.
3. The explanatory tools in set *T* do not posit mental representations.

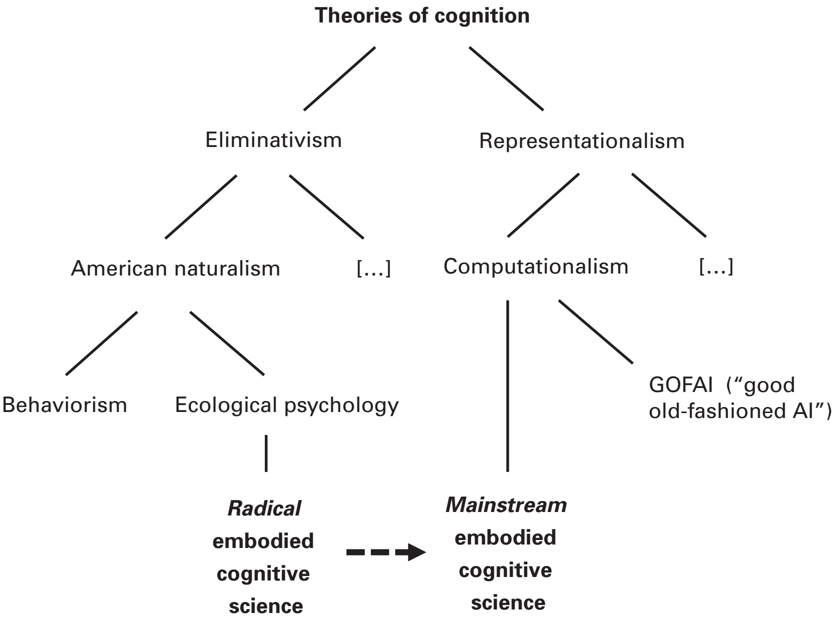

Figure 11.1
Current notions of embodied cognitive science and their historical roots. *Source:* Adapted from Chemero 2009 and Ziemke 2016.

It might be worth noting, though, that Chemero's fundamental and rather strict distinction between representational and antirepresentational approaches to embodied cognition, while offering one useful perspective, should not necessarily be taken to provide some kind of ground truth. First, that seemingly clear-cut distinction obviously hinges on the assumption that there is a more or less wide agreement on what exactly constitutes a representation. This is not necessarily the case (Haselager et al. 2003; Svensson and Ziemke 2005). Hence, it is not difficult to find embodied AI researchers who reject representation in one paper but argue for the grounding of representations in another (e.g., Brooks 1990, 1991). Second, it could be argued that the outright rejection of representation risks throwing out the baby with the bathwater (to use a somewhat dramatic bodily metaphor). That means that while there might very well be good reason to reject the traditional notion of representation, it might be too early, or simply misguided, to reject the notion of representation altogether. Bickhard (1993, 2009), for example, has strongly criticized the traditional notion of representation, which he refers to as *encodingism*, but has developed an interactivist notion of representation, which is much in line with Gibsonian ecological psychology and other elements of radical embodied cognitive science. Similarly, somatic theories of emotion and consciousness, such as the work of Damasio and Panksepp, constitute "a multi-tiered affectively embodied view of mind" (Panksepp 2005, 63), in which representation does play a central role. In this case, however, it is the brain that is considered to "represent" bodily activity, rather than an agent that holds an internal representation of its external environment (cf. Ziemke 2016).

While the controversial issue of representation is certainly too complex to resolve in this chapter, it here suffices to say that, although much early work on embodied cognition (e.g., Varela et al. 1991) and embodied AI (e.g., Brooks 1991; Beer 1995) was explicitly antirepresentationalist, nowadays much of mainstream embodied cognitive science is more or less representationalist. However, there are at least some accounts of embodied cognition that reject the functionalism/computationalism that characterizes the approaches on the right side of Chemero's diagram without, however, rejecting altogether the notion of representation. Hence, as illustrated in figure 11.2, a revised version of Chemero's diagram, more directly relevant in the context of embodied AI and cognitive robotics, could instead be based on the following distinctions:

- On the one hand, some approaches view cognition as (a) embodied and (b) first and foremost a biological phenomenon; some of these are representationalist in some nontraditional sense (e.g., Damasio, Bickhard), and some of them are antirepresentationalist (where the latter roughly correspond to what Chemero referred to as radical embodied cognitive science).
- On the other hand, there are functionalist approaches—such as traditional cognitive science and GOFAI—that view cognition as first and foremost a computational (and representational) phenomenon; among these we can distinguish between the *computational functionalism* of GOFAI and the *robotic functionalism* (Harnad 1989) that characterizes much of Chemero's notion of mainstream embodied cognitive science (according to which cognition is computational, but its representations need to be grounded in sensorimotor interaction with the environment; cf. Harnad 1990; Pezzulo et al. 2013).

Needless to say, the diagram in figure 11.2 should not necessarily be considered as some kind of ground truth either: First, the picture is not complete (behaviorism, for example, is not included). Second, conceptions of representation, computation, embodiment, and so on

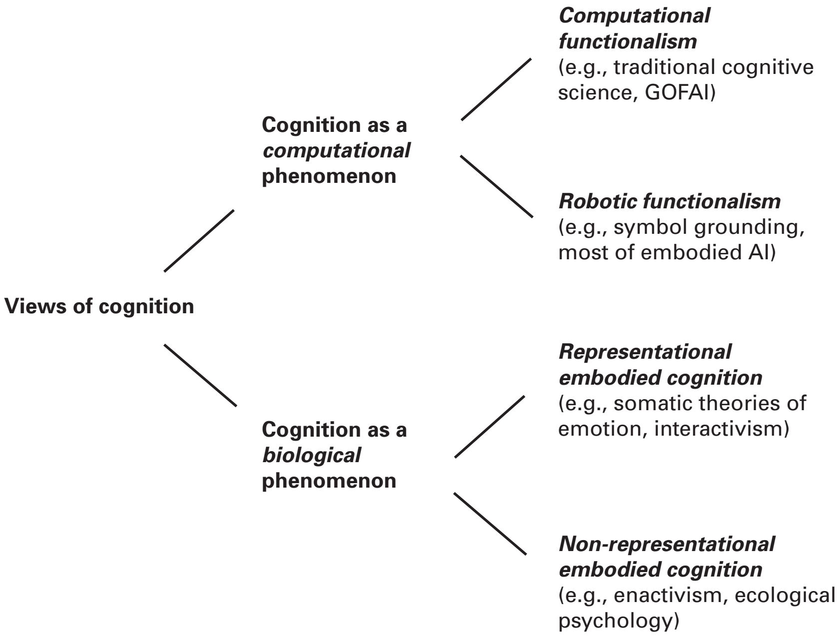

Figure 11.2
Current views of cognition.

obviously vary significantly among researchers. Hence, the reality of the cognitive science and AI research landscape is significantly more complex than either of these diagrams.

11.4 Embodiment in AI

As should be clear from the discussions in previous sections, the issue of embodiment in AI is not straightforward. Many embodied AI researchers, like Brooks (1990) and Steels (1994), emphasize the importance of physical grounding and therefore advocate robotic AI. Others, like Franklin (1997), argue that software agents could be embodied as well if they are situated in an environment (e.g., a search bot searching the internet) and structurally coupled to it. Moreover, many cognitive roboticists in their research practice commonly make use of software simulations of robots and their environments—for example, in order to more quickly train a computational cognitive model in simulation first, which is then later tested on the physical robot. In these cases the computer programs controlling the robots—physical or simulated—are of course for the most part still just as computational as the computer programs of traditional AI. Another case that does not neatly fit into the theoretical categories discussed above are *virtual agents*, as they might appear in video games, for example, or in particular *embodied conversational agents*. Such systems, typically appearing on computer screens, usually have a (simulated) body, used to communicate with their human users, but they typically do not actually use those bodies for sensorimotor interaction with their environment. Hence, while they might *appear* embodied to the people interacting with such systems, in some sense, or to some degree, they really are not embodied in any strong sense.

In fact, most research in embodied AI, although initially often driven by rejections of GOFAI and/or the traditional notion of representation, has been relatively pragmatic in developing the practice of embodied AI, without much concern for philosophical or theoretical distinctions (cf. Ziemke 2004). Based on many years of experience in building embodied AI systems, Pfeifer and colleagues (Pfeifer and Gomez 2005; cf. Pfeifer et al. 2005; Pfeifer and Bongard 2007; Froese and Ziemke 2009) have formulated a number of embodied AI design principles, which together can serve as a characterization of embodied AI as a research field. The first of these are five design procedure principles:

- P1—*synthetic methodology*: aiming for understanding by building.
- P2—*emergence*: systems designed for emergence are often more adaptive.
- P3—*diversity-compliance*: there is a trade-off between exploiting the givens and generating diversity.
- P4—*time perspectives*: three perspectives are required to understand a system's behavior: the "here and now," its ontogeny (development), and its phylogeny (evolution).
- P5—*frame of reference*: the need to distinguish between observed behavior and underlying mechanisms.

These are complemented by eight agent design principles:

- A1—*three constituents*: an agent, its task, and its ecological niche.
- A2—*complete agents*: focus on embodied, situated, autonomous agents.

- A3—*parallel processes*: asynchronous processes, loosely coupled via the environment.
- A4—*sensorimotor coordination*: self-structured/-generated sensory input.
- A5—*cheap design*: systems exploit their niche and interactions.
- A6—*redundancy*: robustness through overlapping functionalities.
- A7—*ecological balance*: between internal and external complexity.
- A8—*value*: systems have driving forces, development, self-organization.

These five plus eight principles can be seen as guidelines for how to design, build, and/or understand embodied AI systems—where the term "embodied" mainly refers to some form of robotic embodiment and the sensorimotor interaction of internal control and external environment over time. As an elaboration of the *value principle* (A8)—which can be questioned in the case of typical robots that might be argued have no own needs or values, given that they do not have bodies that need to self-maintain, survive, and so on—Froese and Ziemke (2009) have formulated two additional *enactive AI* design principles:

1) The system must be capable of generating/maintaining its own systemic identity at some level of description.

2) The system must have the capacity to actively regulate its sensorimotor interaction in relation to viability constraints.

The difference or complementarity between embodied and enactive AI principles can be understood in relation to the theoretical distinctions made in previous sections. While the embodied AI principles of Pfeifer and colleagues mainly emphasize physical/sensorimotor embodiment and structural coupling through sensors and actuators, the enactive AI principles additionally emphasize that the organismic embodiment of living systems implies additional constraints and requirements, but also opportunities, that arise from the fact that living bodies need to self-regulate their internal processes and external interactions so as to remain viable, which implies some kind of bodily homeostasis or allostasis (Froese and Ziemke 2009; cf. Damasio and Carvalho 2013; Vernon et al. 2015; Ziemke 2016).

11.5 Embodiment in Cognitive Robotics

Since this chapter is part of a book on cognitive robotics, we will not dwell on trying to define cognitive robotics in detail (see chapter 1 for a more detailed discussion). For the discussion of embodiment, it might be useful, however, to distinguish roughly between the scientific approach and the engineering approach, although in practice they can certainly overlap.

The engineering approach to cognitive robotics could be characterized as the general endeavor to provide robots with cognitive capacities, such as perception, memory, learning, or communication. An example of this is recent work in our lab, and several others, in the European project DREAM (https://dream2020.github.io/DREAM/), whose aim it was to develop humanoid robots that could interact with kids with autism as part of psychological therapy, with the goal of teaching social interaction skills, such as joint attention, turn-taking, and imitation (e.g., Cao et al. 2019). This is an example of an engineering approach because the mechanisms underlying the robots' cognitive and interactive capacities, for the most part, were not *based on* models of human cognition, although they were of course tailor-made to match the cognitive and interactive capacities of the children involved. Hence, the role of the robot's physical embodiment, much like in the case of the embodied conversational agents mentioned above, is not so much that it is fundamental to the robot's cognitive processes as such but rather that the embodiment plays a crucial role in the kids' embodied social interaction with the robot. For example, the robots needed to be able to perceive the same objects, to observe the kids' behavior, to act (e.g., point to objects), and to communicate (e.g., talk) in ways that the kids could understand.

The (cognitive-) scientific approach to cognitive robotics, on the other hand, could be characterized as the use of robotic models for the express purpose of understanding the mechanisms underlying the cognitive and behavioral capacities of humans and/or other animals. Hence, the contribution of cognitive robotics to the study of embodied cognition lies in building robotic models that help elucidate the many ways in which "cognitive processes are deeply rooted in the body's interactions with the world"—to get back to M. Wilson's (2002) characterization of embodied cognition that we used in the introduction. We can further roughly distinguish between minimalist approaches, which usually try to model general mechanisms or principles underlying cognition and behavior, and more complex approaches, which usually try to build specific models, in many cases aiming to replicate data observed in human or animal experiments.

Some of our own work in evolutionary robotics can be used to illustrate the minimalist approach: Ziemke and Thieme (2002), for example, presented experiments using an evolutionary-robotics methodology (Nolfi and Floreano 2000, chap. 4). Simple simulated wheeled robots were evolved to deal with delayed-response tasks that required "memory" of where light sources had previously been perceived in order to find a goal location in a maze. Delayed response tasks are the classical paradigm in psychology for studying working memory in humans and other animals (Malloy 2011). The point of the simple robotic model in this case was not to replicate the body or the data from some specific animal experiment. It was to illustrate how the embodied cognitive capacity (memory) required to solve the delayed response task (i.e., to "remember" where the light was perceived) could emerge from the interplay of the robot's minimal internal mechanisms and its sensorimotor interaction with its environment, rather than from some explicit internal representation in the traditional sense. The "embodiment" of the robotic agents used in such experiments is often intentionally reduced to a bare minimum, which makes it easy to analyze in detail the interaction of internal and external mechanisms (cf. Ziemke 2003a, 2005), which is of course not possible to do in equivalent experiments with animals. Another example of interacting and coadapting agents would be our work (Buason et al. 2005) on evolving robotic agents in a predator-prey scenario. Here predators and prey were given the opportunity to coevolve—that is, adapt to each other, over a series of generations. In a nutshell, the results showed several effects that have also been observed in natural predator-prey coevolution, such as the fact that predators tended to evolve a narrow field of view (suited to pursue the prey in front of them), whereas the prey evolved a significantly wider field of view (suited to detect both obstacles in front of them and predators behind them). Like the above delayed response tasks, this is another example of minimally cognitive behavior (Beer 1996; Barandiaran and Moreno 2006) because predators and prey had to “remember” each other whenever they temporarily lost track of each other. Again, the “embodiment” in these simulation experiments was intentionally kept minimal—a number of sensors on a simulated, simple circular robot body—and the experiments did not replicate data from some *specific* experiment or species but rather provided insight into the *general* mechanisms of predator-prey coadaptation of body and behavior.

An example of a more complex cognitive robotics experiment aiming to model specific aspects of human cognition, and to also replicate human experimental data, comes from the work of Morse et al. (2015). Using a full-scale humanoid robot, they replicated infant studies investigating the role of bodily posture in how infants learn mappings between words and objects (see chapter 20). The robot model was used to test the hypothesis that a body-centric spatial location, and thus its momentary posture, is used to bind the multimodal features of visual objects and their names. The robot model was shown to replicate data from infant studies and generate novel predictions, which were then tested in new infant studies. This model showed how the memory of name-object mappings, used in new spatial locations, can emerge through the body’s momentary disposition in space.

Hence, although the above robotic models differ radically in the complexity of the robotic embodiment used (simple simulated wheeled robots vs. a physical humanoid), they all address how a cognitive capacity such as memory can emerge—in an embodied manner—from an agent’s sensorimotor interactions with its environment over time.

A negative take on the somewhat perplexing diversity of “embodiments” used by researchers in embodied AI and cognitive robotics would be that 1) researchers simply have not come to any significant agreement on what embodiment “is,” and/or 2) all existing embodied AI systems are still at best very limited versions of the real thing—that is, human or other living bodies. Dreyfus (2007), for example, argued the latter point in his critique of embodied AI, stating that attempts to model human cognition are more or less doomed because they would require “a detailed description of our body and motivations like ours” and that such models “haven’t a chance of being realized in the real world.” However, as we have argued in more detail elsewhere (Froese and Ziemke 2009), the purpose of a model, of course, is usually not to replicate or (re-) instantiate a particular phenomenon in its entirety but rather to help explain it (cf. Morse and Ziemke 2008; Di Paolo and Iizuka 2008). Or, as Froese and Ziemke (2009, 470–471) put it, “It could also be argued that such a detailed modeling approach is not even desirable in the first place since it does not help us to understand why having a particular body allows things in the environment to show up as significant for the agent possessing that body. . . . In other words, instead of blindly modeling the bodies of living beings in as much detail and complexity as possible, it would certainly be preferable to determine the necessary conditions for the constitution of an individual agent with a meaningful perspective on the world.”

As the philosophically minded reader might have noticed by now, this discussion is of course closely related to Searle’s (1980) classical distinction between what he called “weak AI,” the claim that computational models can contribute to our scientific understanding of human cognition, and what he referred to as “strong AI,” the claim that computer models can actually constitute or replicate human or humanlike cognition, consciousness, and so on. Or, in the terms of the discussion in previous sections, if you find yourself sympathizing with functionalist/computationalist approaches to embodied cognition, then you are likely to think that cognitive robotics could lead to a robotic "strong AI," at least in theory. The functionalist position has been formulated explicitly by Zlatev (2001, 155), who posited that "a robot with bodily structures, interaction patterns and development similar to those of human beings . . . could possibly recapitulate [human] ontogenesis, leading to the emergence of intentionality, consciousness and meaning." If, on the other hand, you consider cognition first and foremost a biological phenomenon, then you are likely to think that cognitive robotics is limited to being a form of "weak AI" in Searle's sense—that is, an approach to the scientific modeling of embodied cognition rather than striving for replication of human or humanlike cognition, in a strong sense. While the term "weak" might sound negative, it should be noted that it is only weak in the sense that it is the weaker—or the more realistic, some would say—of the two claims (weak vs. strong AI). As the examples of cognitive robotics models discussed in this section illustrate, this approach can certainly make strong contributions to our scientific understanding of the mechanisms underlying embodied cognition, especially when used as a complement to other scientific approaches and methodologies.

11.6 Conclusion

We started this chapter by characterizing research on embodiment as guided by the view that "cognitive processes are deeply rooted in the body's interactions with the world" (M. Wilson 2002, 625). In the context of AI and robotics, it is then natural to ask what kind of body an artifact with certain cognitive capacities might require and how such a system might become grounded (cf. Harnad 1990; Ziemke 1999) in its environment in roughly the way that humans and other animals are. As the discussion in section 11.2 showed, there is not much agreement regarding the first question: What kind of body? The answers range from software agents, to physical robots, to living organisms. In section 11.3 we discussed that we might need to distinguish between at least two fundamentally different conceptions of embodied cognition, which Chemero (2009) referred to as mainstream embodied cognitive science, which has inherited the representationalism and computationalism of traditional cognitive science, and radical embodied cognitive science, which rejects representationalism—at least according to Chemero (2009). We suggested that a more relevant distinction might be between the functionalist view of cognition as first and foremost a computational phenomenon (which at least in principle should be implementable in robots) and the antifunctionalist view of cognition as first and foremost a biological phenomenon (which simply might not be replicable, at least not with current robotic and computational technologies). As discussed in section 11.4, much—if not most—research in embodied AI is somewhat indifferent to such theoretical distinctions (cf. Ziemke 2004), although antirepresentationalism has long been a driving force in early embodied AI. Instead, much embodied AI research has been driven more by the development of a practice of embodied AI—that is, how should we synthesize and analyze embodied forms of (artificial) intelligence—here characterized with a number of embodied and enactive AI design principles. Section 11.5 then discussed the role of embodiment in cognitive robotics and illustrated this with examples of both minimalistic and more complex/human-level robotic models of embodied cognition, in particular how memory can emerge from embodied agent-environment interactions.

From an engineering perspective, the diversity of concepts and approaches discussed here might be somewhat disappointing—we still don't know how to build *Westworld*-type humanlike robots or even *if* we will ever be able to. However, from the scientific perspective of cognitive robotics as an approach to modeling human and animal cognition, the diversity of approaches is actually rather promising because different approaches can complement each other. As the above examples illustrate, cognitive robotics models can function as both

- a complement to theoretical discussions (e.g., helping to clarify overly abstract discussions of "representation" by concrete models of possible underlying mechanisms) and as
- a complement to empirical studies of cognition in humans and animals (e.g., offering better opportunities for the replication and analysis of experiments, as well as a mechanism of hypothesis testing and generation).

A useful step forward for future research on embodied cognition and cognitive robotics, though, might be a clearer distinction between the different "bodies" or perspectives being addressed. From the discussions in this chapter, we can conclude that we have to distinguish between at least

- the *social body*, as it appears to others,
- the *sensorimotor body*, which interacts with the environment,
- the *living body*, which has to self-regulate and self-maintain, and
- the *lived body*, as it is experienced by an agent itself.

In living systems, these four "bodies" are of course really different aspects of the same body, and they roughly correspond to the overlapping perspectives of different disciplines, such as social psychology, behavioral psychology/ethology, biology, and phenomenology, respectively. These multiple bodies also roughly correspond to the first-person (lived body), second-person (social body), and third-person (living, sensorimotor body) perspectives that are fundamental to much of our social cognition and language use.

Moreover, in living systems these multiple bodies are nested in some sense like Russian dolls, to use one final bodily metaphor; that is, the living body motivates and regulates the sensorimotor body's interaction with the environment, which in turn facilitates and manifests the social body and its interactions. In embodied AI and cognitive robotics, however, some of those Russian dolls are usually missing: Most robots have physical/sensorimotor bodies that are not driven by the needs and motivations of an underlying living body. Furthermore, an artificial agent—most obvious in the case of many embodied conversational agents—might appear to have a social body, although it is not necessarily driven and grounded by a sensorimotor body.

While for researchers in cognitive robotics all of this might be relatively transparent, it remains to be seen exactly how this affects the public perception of robotic systems with cognitive and interactive capacities (e.g., Thellman and Ziemke 2020, 2021)—and in particular how it affects people's embodied social interaction with such diversely embodied technologies as humanoid robots, virtual agents, and automated vehicles (cf. Ziemke 2020). Some of those Russian dolls might not be easy to unpack.

Additional Reading and Resources

- *The* classical book on the embodied mind and the starting point for enactive cognitive science: Varela, F. J., E. Thompson, and E. Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- A broad introduction and comprehensive overview of the research area: Shapiro, L. 2010. *Embodied Cognition*. London: Routledge.
- A review paper with a focus on embodied cognition as a biological phenomenon: Ziemke, T. 2016. "The Body of Knowledge: On the Role of the Living Body in Grounding Embodied Cognition." *BioSystems* 148:4–11.

References

Barandiaran, Xabier, and Alvaro Moreno. 2006. "On What Makes Certain Dynamical Systems Cognitive: A Minimally Cognitive Organization Program." *Adaptive Behavior* 14 (2): 171–185.

Beer, Randall D. 1995. "A Dynamical Systems Perspective on Agent-Environment Interaction." *Artificial Intelligence* 72 (1–2): 173–215.

Beer, Randall D. 1996. "Toward the Evolution of Dynamical Neural Networks for Minimally Cognitive Behavior." In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, edited by P. Maes, M. Mataric, J. A. Meyer, J. Pollack, and S. Wilson, 421–429. Cambridge, MA: MIT Press.

Bickhard, Mark H. 1993. "Representational Content in Humans and Machines." *Journal of Experimental and Theoretical Artificial Intelligence* 5 (4): 285–333.

Bickhard, Mark H. 2009. "The Biological Foundations of Cognitive Science." *New Ideas in Psychology* 27 (1): 75–84.

Black, Daniel. 2014. *Embodiment and Mechanisation: Reciprocal Understandings of Body and Machine from the Renaissance to the Present*. Farnham, UK: Ashgate.

Brooks, R. A. 1990. "Elephants Don't Play Chess." *Robotics and Autonomous Systems*, no. 6: 1–2.

Brooks, Rodney A. 1991. "Intelligence without Representation." *Artificial Intelligence* 47 (1–3): 139–159.

Buason, Gunnar, Nicklas Bergfeldt, and Tom Ziemke. 2005. "Brains, Bodies, and Beyond: Competitive Co-evolution of Robot Controllers, Morphologies and Environments." *Genetic Programming and Evolvable Machines* 6 (1): 25–51.

Cao, Hoang-Long, Pablo G. Esteban, Madeleine Bartlett, Paul Baxter, Tony Belpaeme, Erik Billing, Haibin Cai, et al. 2019. "Robot-Enhanced Therapy: Development and Validation of Supervised Autonomous Robotic System for Autism Spectrum Disorders Therapy." *IEEE Robotics and Automation Magazine* 26 (2): 49–58.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. MIT Press.

Clark, Andy. 1997. *Being There*. Cambridge, MA: MIT Press.

Clark, Andy. 1999. "An Embodied Cognitive Science?" *Trends in Cognitive Science* 3 (9): 345–351.

Damasio, Antonio R. 1998. "Emotion in the Perspective of an Integrated Nervous System." *Brain Research Reviews* 26 (2–3): 83–86.

Damasio, Antonio R. 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Boston: Houghton Mifflin Harcourt.

Damasio, Antonio, and Gil B. Carvalho. 2013. "The Nature of Feelings: Evolutionary and Neurobiological Origins." *Nature Reviews Neuroscience* 14 (2): 143–152.

Di Paolo, Ezequiel A., and Hiroyuki Iizuka. 2008. "How (Not) to Model Autonomous Behavior." *Biosystems* 91 (2): 409–423.

Dreyfus, Hubert L. 1979. *What Computers Can't Do*. Cambridge, MA: MIT Press.

Dreyfus, Hubert L. 2007. "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian." *Philosophical Psychology* 20 (2): 247–268.

Franklin, Stan. 1995. *Artificial Minds*. Cambridge, MA: MIT Press.

Franklin, Stan. 1997. "Autonomous Agents as Embodied AI." *Cybernetics and Systems* 28 (6): 499–520.

Froese, Tom, and Tom Ziemke. 2009. “Enactive Artificial Intelligence: Investigating the Systemic Organization of Life and Mind.” *Artificial Intelligence* 173 (3–4): 466–500.

Gallagher, Shaun. 2005. *How the Body Shapes the Mind*. Oxford: Oxford University Press.

Gallese, Vittorio. 2005. “Embodied Simulation: From Neurons to Phenomenal Experience.” *Phenomenology and the Cognitive Sciences* 4 (1): 23–48.

Gallese, Vittorio, and George Lakoff. 2005. “The Brain’s Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge.” *Cognitive Neuropsychology* 22 (3–4): 455–479.

Goldinger, Stephen D., Megan H. Papesh, Anthony S. Barnhart, Whitney A. Hansen, and Michael C. Hout. 2016. “The Poverty of Embodied Cognition.” *Psychonomic Bulletin and Review* 23 (4): 959–978.

Harnad, Stevan. 1989. “Minds, Machines and Searle.” *Journal of Experimental and Theoretical Artificial Intelligence* 1 (1): 5–25.

Harnad, Stevan. 1990. “The Symbol Grounding Problem.” *Physica D: Nonlinear Phenomena* 42 (1–3): 335–346.

Haselager, Pim, André De Groot, and Hans van Rappard. 2003. “Representationalism vs. Anti-representationalism: A Debate for the Sake of Appearance.” *Philosophical Psychology* 16 (1): 5–24.

Hutchins, Edwin. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.

Johnson, Mark. 2007. *The Meaning of the Body: Aesthetics of Human Understanding*. Chicago: University of Chicago Press.

Kirsh, David, and Paul Maglio. 1994. “On Distinguishing Epistemic from Pragmatic Action.” *Cognitive Science* 18 (4): 513–549.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, George, and Mark Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lindblom, Jessica. 2015. *Embodied Social Cognition*. Vol. 26. Berlin: Springer.

Lund, Henrik Hautop, Barbara Webb, and John Hallam. 1998. “Physical and Temporal Scaling Considerations in a Robot Model of Cricket Calling Song Preference.” *Artificial Life* 4 (1): 95–107.

Malloy, Paul. 2011. “Delayed Response Tasks.” In *Encyclopedia of Clinical Neuropsychology*, edited by J. S. Kreutzer, J. Deluca, and B. Caplan. New York: Springer.

Maturana, Humberto R., and Francisco J. Varela. 1980. *Autopoiesis and Cognition*. Dordrecht: Reidel.

Maturana, Humberto R., and Francisco J. Varela. 1987. *The Tree of Knowledge—the Biological Roots of Human Understanding*. Boston: Shambhala.

Morse, Anthony F., Viridian L. Benitez, Tony Belpaeme, Angelo Cangelosi, and Linda B. Smith. 2015. “Posture Affects How Robots and Infants Map Words to Objects.” *PLoS One* 10 (3): e0116012.

Morse, Anthony F., Carlos Herrera, Robert Clowes, Alberto Montebelli, and Tom Ziemke. 2011. “The Role of Robotic Modelling in Cognitive Science.” *New Ideas in Psychology* 29 (3): 312–324.

Morse, Anthony F., and Tom Ziemke. 2008. “On the Role(s) of Modelling in Cognitive Science.” *Pragmatics and Cognition* 16 (1): 37–56.

Nolfi, Stefano, and Dario Floreano. 2000. *Evolutionary Robotics*. Cambridge, MA: MIT Press.

Panksepp, Jaak. 2005. “Affective Consciousness: Core Emotional Feelings in Animals and Humans.” *Consciousness and Cognition* 14 (1): 30–80.

Pezzulo, Giovanni, Lawrence W. Barsalou, Angelo Cangelosi, Martin H. Fischer, Ken McRae, and Michael Spivey. 2013. “Computational Grounded Cognition: A New Alliance between Grounded Cognition and Computational Modeling.” *Frontiers in Psychology* 3:612.

Pfeifer, Rolf, and Josh Bongard. 2007. *How the Body Shapes the Way We Think: A New View of Intelligence*. Cambridge, MA: MIT Press.

Pfeifer, Rolf, and Gabriel Gomez. 2005. “Interacting with the Real World: Design Principles for Intelligent Systems.” *Artificial Life and Robotics* 9 (1): 1–6.

Pfeifer, Rolf, Fumiya Iida, and Josh Bongard. 2005. “New Robotics: Design Principles for Intelligent Systems.” *Artificial Life* 11 (1–2): 99–120.

Pfeifer, Rolf, and Christian Scheier. 1999. *Understanding Intelligence*. Cambridge, MA: MIT Press.

Quick, Tom, Kerstin Dautenhahn, Chrystopher L. Nehaniv, and Graham Roberts. 1999. “On Bots and Bacteria: Ontology Independent Embodiment.” In *Proceedings of the 5th European Conference on Advances in Artificial Life (ECAL ’99)*, edited by D. Floreano et al., 339–343. Berlin: Springer-Verlag.

Riegler, Alexander. 2002. “When Is a Cognitive System Embodied?” *Cognitive Systems Research* 3 (3): 339–348.

Searle, John. 1980. “Minds, Brains, and Programs.” *Behavioral and Brain Sciences* 3 (3): 417–457.

Shapiro, Lawrence. 2010. *Embodied Cognition*. London: Routledge.

Steels, Luc. 1994. “The Artificial Life Roots of Artificial Intelligence.” *Artificial Life* 1 (1–2): 75–110.

Svensson, Henrik, and Tom Ziemke. 2005. “Embodied Representations: What Are the Issues?” In *Proceedings of the Annual Meeting of the Cognitive Science Society* (27).

Thellman, Sam, and Tom Ziemke. 2020. “Do You See What I See? Tracking the Perceptual Beliefs of Robots.” *iScience* 23 (10): 101625.

Thellman, Sam, and Tom Ziemke. 2021. “The Perceptual Belief Problem: Why Explainability Is a Tough Challenge in Social Robotics.” *ACM Transactions on Human-Robot Interaction* 10 (3): 29.

Thompson, Evan. 2007. *Mind in Life*. Cambridge, MA: Harvard University Press.

Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Vernon, David, Robert Lowe, Serge Thill, and Tom Ziemke. 2015. “Embodied Cognition and Circular Causality: On the Role of Constitutive Autonomy in the Reciprocal Coupling of Perception and Action.” *Frontiers in Psychology* 6:1660.

Wilson, Andrew D., and Sabrina Golonka. 2013. “Embodied Cognition Is Not What You Think It Is.” *Frontiers in Psychology* 4:58.

Wilson, Margaret. 2002. “Six Views of Embodied Cognition.” *Psychonomic Bulletin and Review* 9 (4): 625–636.

Ziemke, Tom. 1999. “Rethinking Grounding.” In *Understanding Representation in the Cognitive Sciences*, edited by A. Riegler, M. Peschl, and A. Von Stein. New York: Plenum Press.

Ziemke, Tom. 2001. “Are Robots Embodied?” In Vol. 85, *Proceedings of the First International Workshop on Epigenetic Robotics: Modelling Cognitive Development in Robotic Systems*, edited by C. Balkenius, J. Zlatev, C. Brezeal, K. Dautenhahn, and H. Kozima, 75–83. Lund, Sweden: University Cognitive Studies.

Ziemke, Tom. 2003a. “On the Role of Robot Simulations in Embodied Cognitive Science.” *AISB Journal* 1 (4): 389–399.

Ziemke, Tom. 2003b. “What’s That Thing Called Embodiment?” In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, edited by R. Alterman and D. Kirsh, 1305–1310. Mahwah, NJ: Lawrence Erlbaum.

Ziemke, Tom. 2004. “Embodied AI as Science: Models of Embodied Cognition, Embodied Models of Cognition, Or Both?” In *Embodied Artificial Intelligence*, edited by F. Iida, R. Pfeifer, L. Steels, and Y. Kuniyoshi, 27–36. Heidelberg: Springer.

Ziemke, Tom. 2005. “Cybernetics and Embodied Cognition: On the Construction of Realities in Organisms and Robots.” *Kybernetes* 34 (1/2): 118–128.

Ziemke, Tom. 2016. “The Body of Knowledge: On the Role of the Living Body in Grounding Embodied Cognition.” *Biosystems* 148:4–11.

Ziemke, Tom. 2020. “Understanding Robots.” *Science Robotics* 5 (46): eabe2987.

Ziemke, Tom, and Mikael Thieme. 2002. “Neuromodulation of Reactive Sensorimotor Mappings as a Short-Term Memory Mechanism in Delayed Response Tasks.” *Adaptive Behavior* 10 (3–4): 185–199.

Ziemke, Tom, and Serge Thill. 2014. “Robots Are Not Embodied! Conceptions of Embodiment and Their Implications for Social Human-Robot Interaction.” In *Robophilosophy*, edited by Johanna Seibt, Raul Hakli, and Marco Nørskov, 49–53. Amsterdam: IOS Press.

Ziemke, Tom, Jordan Zlatev, and Roslyn M. Frank. 2006. *Body, Language and Mind. Vol. 1: Embodiment*. Berlin: Mouton De Gruyter.

Zlatev, Jordan. 2001. “The Epigenesis of Meaning in Human Beings, and Possibly in Robots.” *Minds and Machines* 11 (2): 155–195.

Cognitive Robotics

Edited by: Angelo Cangelosi, Minoru Asada

Citation:

Cognitive Robotics
Edited by: Angelo Cangelosi, Minoru Asada
DOI: 10.7551/mitpress/13780.001.0001
ISBN (electronic): 9780262369329
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Times New Roman by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Cangelosi, Angelo, 1967– editor. | Asada, Minoru, editor.
Title: Cognitive robotics / edited by Angelo Cangelosi and Minoru Asada.
Other titles: Cognitive robotics (M.I.T. Press)
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series: Intelligent robotics and
autonomous agents series | Includes bibliographical references and index.
Identifiers: LCCN 2021031320 | ISBN 9780262046831 (hardcover)
Subjects: LCSH: Autonomous robots.
Classification: LCC TJ211.35 .C628 2022 | DDC 629.8/92—dc23
LC record available at https://lccn.loc.gov/2021031320