BioSystems 148 (2016) 4–11

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

Review article

# The body of knowledge: On the role of the living body in grounding embodied cognition

Tom Ziemke[a,b,*]

[a] Interaction Lab, School of Informatics, University of Skövde, 54128 Skövde, Sweden
[b] Cognition & Interaction Lab, Human-Centered Systems Division, Department of Computer & Information Science, Linköping University, 58183 Linköping, Sweden

ARTICLE INFO

*Article history:*
Received 6 August 2016
Accepted 9 August 2016
Available online 16 August 2016

*Keywords:*
Allostasis
Cognitive systems
Grounding
Homeostasis
Embodied AI
Embodied cognition
Emotion
Intentionality
Predictive regulation
Representation

ABSTRACT

Embodied cognition is a hot topic in both cognitive science and AI, despite the fact that there still is relatively little consensus regarding what exactly constitutes 'embodiment'. While most embodied AI and cognitive robotics research views the body as the physical/sensorimotor interface that allows to ground computational cognitive processes in sensorimotor interactions with the environment, more biologically-based notions of embodied cognition emphasize the fundamental role that the living body – and more specifically its homeostatic/allostatic self-regulation – plays in grounding both sensorimotor interactions and embodied cognitive processes. Adopting the latter position – a multi-tiered affectively embodied view of cognition in living systems – it is further argued that modeling organisms as layered networks of bodily self-regulation mechanisms can make significant contributions to our scientific understanding of embodied cognition.



**Contents**



## 1. Introduction

At some point in the long and winding write-up of this paper, its title was *"What makes embodied AI embodied?"*. That title eventually disappeared again, but the question is still highly relevant to this paper – and the answer is not as straightforward as one might think. Robots are, no doubt, considered 'embodied' by most AI researchers, and in fact the obvious AI approach to modeling natural embodied cognition or synthesizing artificial equivalents thereof (cf., e.g. Ziemke, 2003; Morse et al., 2011). Much embodied AI research, however, also makes use of simulated robots or other types of non-physical agents, e.g. so-called virtual agents or different types of more abstract artificial-life agents. Hence, one might ask (cf. Ziemke, 2004) whether embodied AI really is about embodied (i.e. physical, robotic, etc.) models of cognition, or rather about models – any type of model: robotic ones obviously, but also purely computational ones – of embodied cognition (whatever that is), or maybe both? If you find that question somewhat confusing, you are not alone. As discussed in more detail in Section 2, despite more than 25 years of research on embodied cognition and AI, and by now a number of books on the topic (e.g. Varela et al., 1991; Clark, 1997; Pfeifer and Scheier, 1999; Gallagher, 2005; Ziemke et al., 2006; Johnson, 2007; Thompson, 2007; Shapiro, 2010; Lindblom,

* Correspondence address: IDA, Linköping University, 58183 Linköping, Sweden.
*E-mail addresses:* tom.ziemke@liu.se, tom.ziemke@his.se

http://dx.doi.org/10.1016/j.biosystems.2016.08.005

2015), there still is a perplexing diversity of notions of embodied cognition as well as claims concerning its nature and relevance.

Given that this paper is part of a journal special issue on the relation between embodied AI and synthetic biology, it should come as no surprise that it is argued here *that* synthetic biology research might be able to make significant contributions to embodied AI (the details of *how* are beyond the scope of this paper though) – and thereby also might help to clarify the role that biological embodiment plays in natural cognition. To what degree the underlying biological mechanisms really do play a role in cognitive processes and capacities, is another open question in the cognitive sciences, and in fact not everybody would agree that they actually do play any role at all, other than that of a particular physical implementation that could just as well be replaced by another, non-biological – e.g. computational and/or robotic – implementation. Different arguments supporting the view that the underlying biology in general, and bodily self-regulation in particular, actually does play a crucial role in embodied cognition are discussed in more detail in Section 3.

Section 4 then, finally, presents some discussion and conclusions. It will be argued that embodied cognition is not only grounded in sensorimotor interaction with the environment – a claim that most proponents of embodied cognition, and even some of its opponents, would agree to – but that at least natural cognition is furthermore also deeply rooted in the underlying biological mechanisms, and more specifically layered/nested networks of homeostatic/allostatic bodily self-regulation mechanisms. Hence, the potential contribution of synthetic biology to embodied cognition and AI, it will be argued, lies first and foremost in modeling/understanding/synthesizing the nature of organisms as such layered networks. This would be an important complement to current work in embodied AI and cognitive architectures/robotics, much of which is predominantly concerned with layered architectures for dealing with the complexities of perceiving and acting in the external environment.

## 2. What's that thing called embodiment?

The embodied approach in cognitive science and AI has received increasingly much attention in recent years. In fact, *"Embodied Cognition is sweeping the planet"*, at least according to Fred Adams' backcover book endorsement of the paperback edition of Shapiro's (2010) book on the topic. Research on embodied cognition has received significant attention in the cognitive sciences for at least 25 years now, if you count from the appearance of Varela, Thompson and Rosch's book *"The Embodied Mind"* in 1991. It should be noted though that despite this, at least at this point in time, there actually is no such thing as *the* embodied mind thesis or paradigm. This is reflected, for example, by recent paper titles such as *"Embodied cognition is not what you think it is"* (Wilson and Golonka, 2013) and recent debates about the alleged "poverty of embodied cognition" (Goldinger et al., 2016; Killeen, 2016) that reveal deep misunderstandings and wildly different (mis-) conceptions of even the most basic tenets of embodied cognition research.

From the embodied AI researcher's perspective, on the other hand, what is and what is not embodied might seem relatively straightforward: the computer programs of traditional AI research are widely considered 'disembodied', whereas robots obviously are embodied – at least in some sense (cf. Ziemke, 2001b; Ziemke and Thill, 2014). Much early embodied AI research was to some degree driven by criticisms of traditional AI formulated by philosophers such as Dreyfus (1979), Searle (1980) and Harnad (1990). A key point in these criticisms was the lack of interaction between the internal representations – at the time typically symbolic ones – of AI programs and the external world they were supposed to represent. Dreyfus (1979), for example, inspired by Heidegger's notion of being-in-the-world, argued that any computer program "is not always-already-in-a-situation. Even if it represents all human knowledge in its stereotypes, including all possible types of human situations, it represents them from the outside . . . It isn't situated in any one of them, and it may be impossible to program it to behave as if it were". Searle's (1980) criticism of computational AI systems, based on his famous *Chinese Room Argument*, was that "the operation of such a machine is defined solely in terms of computational processes over formally defined elements", and that such "formal properties are not by themselves constitutive of intentionality" – which is the characteristic of human cognition that allows it to be *about* the world. Harnad's (1990) argument was based on Searle's, but he referred to the problem of intentionality as a lack of 'intrinsic meaning' in purely computational systems, which he argued could be resolved by what he termed *symbol grounding*, i.e. the grounding of internal symbolic representations in sensorimotor interactions with the environment.

*Embodied* approaches to AI – using robotic or simulated 'autonomous agents' – at least at a first glance, allow computer programs and the representations they are using, if any, to be grounded in interactions with the physical environment through the robot/agent platform's sensorimotor capacities. Brooks, for example, one of the pioneers of embodied AI, formulated what he called "the two cornerstones of the new approach to Artificial Intelligence, situatedness and embodiment" (Brooks, 1991). Embodiment from this perspective simply means that "robots have bodies and experience the world directly – their actions are part of a dynamic with the world and have immediate feedback on their own sensations" (Brooks, 1991). According to Brooks, such systems are physically grounded, and hence internally "everything is grounded in primitive sensor motor patterns of activation" (Brooks, 1993). Situatedness, accordingly, means that "robots are situated in the world – they do not deal with abstract descriptions, but with the here and now of the world directly influencing the behavior of the system" (Brooks, 1991).

Hence, from the embodied AI perspective, things might seem relatively uncomplicated: robots are embodied and situated in roughly the same sense that humans and other animals are, and thereby they at least potentially can overcome traditional AI's problems with intentionality or intrinsic meaning. The problem of computer programs dealing with ungrounded representations is solved through physical grounding and either not having any representations at all (a la Brooks) or acquiring internal representations through symbol/representation grounding (a la Harnad), i.e. developing such representations in the course of interaction with the external world (e.g. learning a map of the environment). It should be noted though that this does not necessarily resolve the philosophical problems discussed above. Searle, for example, already back in 1980, presented – and rejected – what he called the *'robot reply'* to his own *Chinese Room Argument*. This entailed pretty much exactly what is now called *embodied AI*, namely computer programs running inside robots that interact with their environment through sensors and actuators. In the terms of Searle's argument, to the person inside the Chinese Room, it does not make any difference whether or not inputs to and outputs from the Chinese Room are connected to the sensors and motors of a robot – the person inside the room still lacks the *intentionality* that characterizes human cognition.

At this point it should be noted that for the purposes of this paper it does not actually matter at all whether or not the reader is familiar with the details of Searle's *Chinese Room Argument*, let alone convinced of its validity. The argument has been discussed for more than 35 years now (e.g. Harnad, 1989, 1990; Ziemke, 1999; Zlatev, 2001; Preston and Bishop, 2002) without reaching much consensus. What is more interesting here though is that there are quite many