



# Human-extended machine cognition

Action editor: George Kampis

Paul R. Smart

*Electronics and Computer Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK*

Received 7 May 2017; accepted 6 November 2017

Available online 21 November 2017

## Abstract

Human-extended machine cognition is a specific form of artificial intelligence in which the causally-active physical vehicles of machine-based cognitive states and processes include one or more human agents. Human-extended machine cognition is thus a specific form of extended cognition that sees human agents as constituent parts of the physical fabric that realizes machine-based cognitive capabilities. This idea is important, not just because of its impact on current philosophical debates about the extended character of human cognition, but also because it helps to focus scientific attention on the potential role of the human social environment in realizing novel forms of artificial intelligence. The present paper provides an overview of human-extended machine cognition and situates the concept within the broader theoretical framework of active externalism. The paper additionally shows how the concept of human-extended machine cognition can be applied to existing forms of human–machine interaction, especially those that occur in the context of the contemporary Internet and Web.

© 2017 Elsevier B.V. All rights reserved.

*Keywords:* Extended cognition; Active externalism; Artificial intelligence; Cognitive computing; Human–computer interaction; Human computation

## 1. Introduction

In the attempt to understand the mechanisms of human intelligence, the human brain is a structure of central significance. In fact, given the way in which cognitive phenomena are seen to emerge from the whirrings and grindings of the neurological machinery, it might be assumed that all that matters to human intelligence (from the standpoint of mechanistic realization, at least) is to be found solely within the neural realm—that the point source of intelligent thought and action is located inside the heads of human agents. This view, which I will dub the *neurocentric view*, sees the biological brain as the sole realization base for the human mind. According to the neurocentric view, human mental states and processes are the direct product

of what the brain does. The human brain, in other words, is that part of the material world that realizes all humanly-relevant cognitive phenomena.<sup>1</sup>

An alternative vision of the mechanistic underpinnings of human intelligence comes in the form of a philosophical position known as active externalism (Clark, 2008; Clark & Chalmers, 1998). In contrast to the neurocentric view, advocates of active externalism propose that the machinery of the human mind is not restricted to the inner sanctum of the neural realm. Instead, they claim that the causally-active physical vehicles of the mind can, on occasion, extend beyond the traditional biological borders of skin

<sup>1</sup> Harnad and Dror (2006) thus embrace the neurocentric view when they suggest that “cognition takes place entirely within the brains of cognizers. . . cognition begins and ends at the cognizer’s sensor and effector surfaces” (p. 9).

E-mail address: [ps02v@ecs.soton.ac.uk](mailto:ps02v@ecs.soton.ac.uk)

and skull to include a range of non-neural (and even non-biological) elements. Active externalism thus provides us with an *extended view* of human cognition—a view in which the physical machinery of the mind is occasionally able to escape its cranial confines and extend out into the world.

Active externalism is a philosophical position of considerable interest and importance to cognitive science. In our quest to understand the material bases of intelligent thought and action, it is obviously important that we focus our attention on those parts of the physical world that are most likely to contain the mechanisms that are responsible for phenomena of interest (Craver & Tabery, 2016). If our attention is focused on only one part of what is, in effect, a larger mechanistically-relevant matrix, then we face the risk that an important array of explanatorily-relevant forces and factors will end up falling beyond our field of view. The result is that a commitment to active externalism implies a shift in scientific focus. For inasmuch as we accept the basic tenets of the active externalist position, then it seems that our attempt to understand human intelligence will need to focus on more than just the biological brain. This looks to be particularly important if it is the porosity of the human cognitive system—i.e., our capacity to assimilate extra-organismic resources deep into our cognitive routines—that lies at the heart of our species' peculiar and prodigious form of cognitive success.<sup>2</sup>

But it is not just our approach to human intelligence that is affected by claims about the extended character of human cognition; active externalism is also a philosophical position that is of substantive relevance to the field of Artificial Intelligence (AI). In one sense, of course, this is trivially true. If the profile of human cognition is one that relies on a capacity for cognitive extension (i.e., a capacity to assimilate extra-organismic elements into cognitive processing routines), then attention to the details of that capacity is likely to yield rewards in terms of the attempt to engineer systems that seek to emulate (or surpass) human capabilities. There is, however, also a sense in which active externalism may be relevant to AI irrespective of whether or not we accept claims about the extended character of human cognition. It may be the case, for example, that an extended perspective helps us see how an array of difficult or intractable computational problems could be solved by exploiting cognitively-potent forms of causal commerce between an array of materially-heterogeneous resources. By combining the distinctive representational and computational capabilities of resources drawn from both the biological and the technological realm it may thus be possible to yield solutions that reduce (e.g.) the temporal and energetic

costs associated with the performance of certain kinds of cognitive task (see Jonker, 2008). In this sense, an extended approach to machine intelligence dovetails with work that emphasizes the value of situated, embodied and enactive approaches to the design of AI systems (Chrisley, 2003; Froese & Ziemke, 2009; Iida, Pfeifer, Steels, & Kuniyoshi, 2004; Lindblom & Ziemke, 2003).

In the present paper, I attempt to extend the conventional focus of active externalist theorizing by outlining an extended approach to AI.<sup>3</sup> In particular, I claim that when we look at the environment in which specific instances of machine intelligence are situated, we sometimes encounter a state-of-affairs in which one or more human agents serve as part of the physical fabric that realizes episodes of machine-based cognizing. This idea is captured by the thesis of Human-Extended Machine Cognition (HEMC):

#### Thesis of Human-Extended Machine Cognition

The casually-active physical vehicles of machine-based cognitive states and processes may, on occasion, include one or more human agents. Human agents are thus candidate parts of the (extended) realization base for machine-based cognitive capabilities.

The HEMC concept, it should be clear, reverses the usual focus of active externalist theorizing. Instead of the idea that (e.g.) a technological device serves as part of the material fabric that realizes a specific instance of human cognizing, the HEMC thesis encourages us to switch our point of view and look at things from the device's 'perspective'. In this case, the nature of the coupling between the device and the human user can be seen to provide opportunities for *bidirectional* forms of cognitive extension. If we accept the idea that a technological artifact may, on occasion, form part of the supervenience base for human cognitive states and processes, then it is surely possible, at least in principle, for the human agent to also, on occasion, form part of the material fabric that realizes the cognitive states and processes of a technological artifact.

The main aim of the present paper is to introduce the HEMC concept and situate it within the broader philosophical and cognitive scientific literature. I also attempt to show how the HEMC concept can be applied to existing forms of human-machine interaction, especially those that occur in the context of the contemporary Internet and Web. Finally, I hope to demonstrate that the HEMC concept enables us to identify a number of important links between (currently) disparate areas of theoretical and empirical research.

<sup>3</sup> The application of active externalist theorizing to AI systems has been the focus of previous work. Jonker (2008), for example, provides a useful overview of some of the potential advantages associated with cognitive extension in the case of AI systems. The focus of the present paper is on a specific form of cognitive extension that involves the incorporation of *human* agents into machine-based cognitive processing routines. This is not something that is explicitly addressed by Jonker.

<sup>2</sup> Clark (2003), for example, notes that “what is special about human brains, and what best explains the distinctive features of human intelligence, is precisely their ability to enter into deep and complex relationships with non-biological, props, and aids” (p. 5).

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات