Cognitive Extension and the Web

Paul Smart[1], Paula Engelbrecht[2], Dave Braines[3], Mike Strub[4] and James Hendler[5]

[1]*School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK.*
[2]*School of Psychology, University of Southampton, Southampton, SO17 1BJ, UK*
[3]*Emerging Technology Services, IBM United Kingdom Ltd, Hursley Park, Winchester, SO21 2JN, UK*
[4]*Army Research Laboratory, Alexandria, Virginia, 22310-3225, USA.*
[5]*Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA.*

Introduction

There has been a growing interest in recent years regarding the relationship between social interaction processes, technological artefacts and the mechanisms that underpin human mental states and processes (Clark, 2003, 2008; Hollan et al., 2000; Hutchins, 1995). According to one view, the external environment in which an agent is embedded constitutes more than just a space for sensory inputs and motor outputs; it is also something that can be flexibly factored into episodes of cognitive processing (Clark, 1997, 2008). Inasmuch as this is true, a perspective that sees all the mechanisms of mind as firmly located 'inside the head' of a particular agent may no longer be appropriate. Instead, in trying to understand the mechanistic underpinnings of the human mind, we may need to place a much greater emphasis on the representational and computational roles played by a complex web of extra-neural resources, resources that include, on occasion, elements of the wider social and technological matrix in which human thought and reason is so often situated.

The view that some parts of the physical machinery associated with human cognition may be externally located is most strongly represented by the 'extended mind hypothesis' (Clark & Chalmers, 1998). If the view is correct then the advent of new technologies, such as the World Wide Web (hereafter referred to as the Web), throws up an interesting question: do such technologies provide opportunities for cognitively-potent forms of bio-technological integration, and, if so, what is the likely impact of such mergers on the space of human thought and reason? What, in other words, is the transformative potential of the Web with respect to the panoply of mental states and processes that we typically identify with the human mind?

The main aim of this paper is to explore the possibility that the technological and informational elements of the Web can (or at least could) constitute part of the material supervenience base for some aspects of the human mind. This evaluation draws on the kind of criteria that have been used by extended mind theorists to defend against an excessive liberalism regarding the environmentally-extended nature of human cognition (Clark & Chalmers, 1998).

The Extended Mind Hypothesis

In the attempt to understand human cognition, cognitive science has tended to focus on the brain as the mechanistic substrate of mental phenomena. Cognitive processes, as well as the familiar elements of mentalistic discourse – the mental states that we typically use to explain and predict human behaviour – seem to be firmly located within the traditional biological boundaries of skin and skull. In contrast to this view, Clark and Chalmers (1998) argue that mental states and processes are not necessarily tied to the biological brain; the physical mechanisms that support these phenomena can also extend into the external environment to incorporate a variety of external technological props, aids and artefacts. This philosophical position is known as 'active externalism', but it is more commonly referred to as the 'extended mind hypothesis'. Extended mind accounts typically focus on the way in which much real-world cognitive processing seems to depend on the exploitation of both bodily contingencies and aspects of the *local* external environment (Clark, 1997). Thus, in solving long multiplication problems, human agents typically resort to using pen and paper to store intermediate solutions as well as structure the overall sequence of information processing steps (Rumelhart et al., 1986). Similarly, in writing an academic paper, the human agent is often engaged in a complex sequence of iterated interactions with a variety of external resources (word processors, books, post-it notes, marginalia, etc.) that, in conjunction with the operation of the human brain, serve to progressively restructure and refine a preliminary set of initial thoughts into a finished article. What these and other examples (Clark, 1997; Hutchins, 1995; Kirsh, 1995; Kirsh & Maglio, 1994) highlight is the importance of external tools, props and artefacts in shaping, augmenting and guiding cognitive processes. The extended mind hypothesis argues that such external resources are not only supportive of cognition, they also form part of the very machinery that makes much of human cognition mechanistically possible.

According to the extended mind hypothesis, extra-neural and extra-organismic resources should, at least on occasion, be seen as playing important roles with respect to the material realization of human cognitive processes. The argument also applies with equal force to human mental states. According to Clark and

Chalmers (1998), external resources can form part of the long-term store of dispositional beliefs and knowledge associated with a human agent, providing that such resources serve to guide sequences of thought and action in a manner that is similar (i.e. functionally equivalent) to information retrieved from long-term bio-memory. Take, as an example, a portable device that is used to store information about future meeting appointments. Inasmuch as the use of this device fulfils a number of conditions (more on which below) we are justified, so claim Clark and Chalmers (1998), in seeing the device as akin to a basic biological resource, such as memory. In providing information, the portable device serves the same causal role in coordinating action that the same information would do were it to be retrieved from long-term bio-memory. As such, the device can be treated as a store of long-term beliefs about the timing and location of future meeting appointments. According to the extended mind hypothesis, the technological accoutrements of the modern age – palm pilots, calculators, mobile phones, etc. – are not simply tools that support our cognitive endeavours, they are also (potentially at least) part of the material fabric that physically realizes human mental states and processes. Inasmuch as this is the case, the transformative potential of new technologies is indeed profound: by developing technologies that can be easily co-opted into the processing loops of an extended cognitive system, engineers are providing opportunities for cognitive extension that promise (or perhaps threaten) to transfigure our traditional notions of cognitive and epistemic (i.e. knowledge-guided) competence.

The extended mind hypothesis has been the subject of an ongoing debate within the philosophy of mind and cognitive science communities. Common criticisms of the extended mind hypothesis include the location of cognitive and computational control within an extended cognitive system (Butler, 1998), the distinction between intrinsic and derived content (Adams & Aizawa, 2001, 2007), and the vulnerability of external resources to damage and social manipulation (Sterelny, 2004). Clark (2008) provides a useful summary of the main arguments both for and against the claims for cognitive extension.

The Web-Extended Mind?

If the claims of the extended mind hypothesis are correct then the machinery of the human mind is not solely located within the human head. What does this potential for extra-corporeal extension mean when it comes to a technology as pervasive and common-place as the Web? Could it be that the Web is poised to participate in the material realization of at least some aspects of our mental states and processes, or is it simply absurd to even contemplate the possibility of a Web-extended mind? In answering this question it is important to understand when an extended mind account is and is not warranted. Clearly, not all of the technologies or external resources that we encounter are apt to engage in the kind of bio-technological hybridization envisioned by the extended mind hypothesis. As Clark (1997) argues "There would be little value in an analysis that credited me with knowing all the facts in the Encyclopaedia Britannica just because I paid the monthly instalments and found space for it my garage" (pg. 217). Similarly, it would be foolish to equate my personal body of knowledge as co-extensive with the contents of the internet simply because I have an internet-enabled mobile phone. What then are the conditions under which we count a set of external resources as participating in genuine cases of cognitive extension?

In order to allay concerns about the excessive liberalism of the extended mind perspective, Clark and Chalmers (1998) offer a set of rough-and-ready criteria that should be met in order to justify claims about cognitive extension. The criteria are:

1. That the resource be reliably and typically invoked [**Availability Criterion**].
2. That any information retrieved from the resource be more or less automatically endorsed. Information should, in general, be deemed about as trustworthy as information retrieved from biological memory [**Trust Criterion**].
3. That information contained in the resource should be easily accessible as and when required [**Accessibility Criterion**].
4. That the information derived from the external resource has been consciously endorsed in the past and indeed is there as a consequence of this endorsement [**Conscious Endorsement Criterion**].

In the sections that follow we briefly evaluate the notion of Web-based cognitive extension with respect to these criteria.

Availability

The first of Clark and Chalmer's criteria (1998) concerns the extent to which an external resource is both available and typically used by a human agent. In order for the informational and technological elements of the Web to feature as part of an extended cognitive system, those elements should be more-or-less constantly available for use. Part of the concern here relates to the portability of the devices used to access information –

the more portable the device, the more likely it is to feature as part of an extended cognitive system. Clearly, the current state-of-the-art in mobile computing features a number of devices that are about as portable as the kind of devices (e.g. paper notebooks) typically encountered in discussions of the extended mind. Well known examples include Personal Digital Assistants (PDAs), Smartphones and ultra-mobile personal computers. In terms of suitability of the Web to support environmentally-extended forms of cognition, therefore, portability does not seem to be a major problem. In fact, mobile computing devices are already in widespread use, and they have largely replaced notebooks and filofaxes as devices that guide daily action and decision-making.

Trust

The second of Clark and Chalmers' (1998) criteria concerns the notion of automatic endorsability or trust. The idea here is that information retrieved from external, non-biological resources should always be trusted to the same extent as information retrieved *clearly* from biological memory. This is an interesting criterion, because it may seem that we can never trust Web-accessible information to the same extent as internal information. Web-derived information exists in a public space, one that provides ample opportunities (and sometimes incentives) for deception and social manipulation. Within such an environment it seems unlikely that we can trust Web-derived information to the same extent as information retrieved from biological memory (see Sterelny, 2004).

One of the reasons for insisting on something like a trust criterion seems to be that the functional poise of information is changed as soon as we see a need to subject it to critical scrutiny. The requirement to engage in epistemic vetting and validation of external information might, for example, prohibit the automatic (i.e. non-conscious) incorporation of that information into ongoing sequences of thought of action. Thus if we have to constantly evaluate the veracity of external information then such information seems ill-suited to guide behaviour in quite the same way as information retrieved from, say, biological memory.

The claim that Web-derived information exists in a public space is undeniable, as is the claim that by accessing such information we are vulnerable to deception and social manipulation. However, whether all this means that Web-based information can never be used to automatically guide behaviour in quite the way we would expect of purely internal (i.e. neurally-derived) information is another matter. As is evidenced by both the psychological phenomenon of 'change blindness' (Simons & Levin, 1997) and the art of the stage magician (Kuhn et al., 2008), we do not always treat externally-derived information with the kind of critical scrutiny and scepticism that it sometimes deserves. Just because people are vulnerable to deception by exploiting Web-based information does not necessarily mean that they actually erect the kind of conscious evaluative mechanisms that might extenuate the possibility of automatic endorsement. In any case, there is now a wealth of research concerning trust and security mechanisms for the Web (e.g. Golbeck, 2008), and it is possible (if not entirely likely) that the technological outcomes of such research will lead to a situation whereby online resources are increasingly trusted and automatically utilized as part of our daily cognitive endeavours.

Accessibility

The third of Clark and Chalmers' (1998) criteria concerns the notion of accessibility. The idea here is that information should be easily and readily accessible, perhaps no less so than the information contained in biological memory. The importance of the accessibility criterion is immediately apparent when we consider what it means for a body of information to count as part of the knowledge (or belief) base of an agent. If we think about the kinds of situations in which we would be prepared to acknowledge that we genuinely know something (e.g. a specific fact) then our ability to access and utilize information (either internally or externally located) seems to be all-important. The longer it takes for us to access information (or the more effortful the information retrieval process becomes), the less likely it seems that such information should count as part of our personal body of knowledge and beliefs. Clearly, the information retrieved from bio-memory often (although perhaps not always) fulfils the accessibility criterion, but what about externally-located information? Clark (2003) presents the case of the humble wristwatch as one example where our access to information may motivate a shift in our intuitions about what counts as part of our personal body of knowledge and beliefs. The advent of the wristwatch has given rise to a situation in which most of us would be inclined to answer in the affirmative when presented with the question 'Do you know the time?' (providing we wear a wristwatch, of course). What seems to count here is not the fact that we always consciously know the time. Rather, the crucial fact is our constant and easy access to a trusted source of temporal information. As Clark (2003) argues, "..it sometimes makes both social and scientific sense to think of your individual knowledge as quite simply whatever body of information and understanding is at your fingertips; whatever body of information and understanding is right there, cheaply and easily available, as and when needed." (pg. 42).

In order for the Web to satisfy the accessibility criterion, therefore, it seems imperative that Web-based information content should be highly accessible. In the ideal case, information should be poised to guide daily sequences of thought and action in a manner that is functionally indistinct from the way in which internally-generated information does. The main problem here is that the kind of devices we typically use to access the Web, as well as the kind of formats in which Web-accessible information content is made available, generally tends to prohibit the kind of rapid access and automatic utilization that we would expect of an extended mind resource. There are clearly some technological advances that might serve to address this gap. They include (*inter alia*) the transition from document-centric to data-centric modes of information encoding (Smart et al., 2008), as well as new forms of Web-directed user interaction and information access[1]. Such advances, in conjunction with improvements to the trust and security infrastructure of the Web, may provide the kind of external information resource that is able to yield behavioural patterns warranting belief ascription.

Conscious Endorsement

As Clark (2008) admits (pg. 80), the importance of the fourth criterion (the conscious endorsement criterion) is somewhat uncertain. On the one hand, such a criterion would seem necessary to avoid excessive liberalism regarding what can and cannot count as part of the supervenience base for states of dispositional belief (we may, for example, want to reject claims that our belief base can be modified by the mere posting of information on the Web by another agent). On the other hand, the conscious endorsement criterion does seem a little too strong. We might be able to acquire beliefs via subliminal perception, implicit learning, and incubation processes, and none of these entails the conscious endorsement of information that might subsequently be used to guide thought and action in adaptive and intelligent ways. Of course, even if we choose to accept the conscious endorsement criterion, this does not mean that the Web cannot act as part of the supervenience base for states of (at least) dispositional belief. It is entirely possible that with the advent of Web 2.0 (in which we see a much greater capacity for user uploading and content personalization) the information contents of at least portions of the web could be consciously endorsed to more or less the same extent as information entered into more local information storage devices (a conventional paper notebook, for example).

Cognition and the Web

Much of our discussion thus far has focused on the possible role of the Web in supporting a specific type of mental state, namely states of dispositional belief. What contribution, if any, does the Web make with regard to the external realization of cognitive *processes*? Given the limited forms of interactivity that are currently possible with many Web resources, we suggest that the potential for cognitive incorporation on the Web is probably weak. When we confront examples of environmentally-extended cognitive processing, whether it be in the area of memory, perception or reasoning (see Rowlands, 2003), what we seem to encounter are cases in which brain, body and world form a closely-coupled system, one in which aspects of the external environment are manipulated so as to influence subsequent internal processing. The case of Scrabble is a case in point (Kirsh, 1995). What we see in the case of Scrabble is an iterative sequence of body-world interactions in which aspects of the external environment (the Scrabble tiles) are manipulated so as to better exploit the computational profile of the human brain, i.e. its capacity for rapid pattern-recognition and pattern-completion. The problem of word search in this case seems to have been partially transformed from an exhaustive search of the space of cryptarithmetic solutions into something more akin to perceptual pattern recognition and completion: candidate or partial word solutions that are encountered in the perceptual array serve to guide tile manipulation actions in ways that facilitate the further derivation of candidate solutions. Expert performance in the game of Tetris has been characterized in similar ways (Kirsh & Maglio, 1994). What these (and other examples) show is that extended cognition often depends on information processing loops whose temporal profiles are geared to work in concert with the fast pattern-recognition and pattern-completing operations of the biological brain. This does not necessarily exclude Web-accessible resources from participating in extended cognitive processes, but it does seem to limit the kind of resources that could feasibly fulfil this role. In general, it seems that only resources supporting a high degree of user interactivity (usually accomplished by client-side code and applets) are apt to extend human cognition in the ways characteristic of more conventional props, aids and artefacts.

[1] Examples include the use of sub-vocalization techniques to support Web navigation (Jorgensen & Binsted, 2005) and the use of mobile device eyewear systems to display information directly to a user's visual field (http://www.microvision.com/wearable_displays/index.html).

Conclusion

The conventional view in cognitive science is that much of the physical machinery of the human mind is located firmly inside the human head. In contrast to this view, the extended mind hypothesis posits that body and world can sometimes form part of the machinery by which mind and cognition are physically realized; body and world hence form part of the local material supervenience base for mental states and processes. In this paper, we have attempted to explore the claim that the technological and informational elements of the Web could, at least in principle, constitute part of the physical machinery associated with a variety of human mental states and processes. We suggest that within the context of the extended mind hypothesis, the Web is most suited to act as an environmentally-extended store of long-term knowledge and dispositional beliefs. While we do not discount the possibility that the Web can participate in cognitive processing functions, the kind of temporal delays associated with current client-server models make it unlikely that all but purely client-side applications could support cases of environmentally-extended cognitive processing. In terms of the external realization of mental states (e.g. states of dispositional belief), we have seen that the Web encounters a number of problems when it comes to satisfying the criteria typically used to identify extended cognitive systems. Perhaps the most important of these problems concerns the functional poise of Web-based information content to guide thought and action in the manner typically associated with cases of purely internal (in-the-head) processing. Another potential problem concerns the inherent credibility and trustworthiness of Web-based information content. This may create barriers to the kind of rapid, effortless and largely automatic mode of information exploitation and utilization that seems to lie at the heart of some of the most compelling cases of environmentally-extended cognition.

References

Adams, F., & Aizawa, K. (2001) The bounds of cognition. *Philosophical Psychology*, 14(1), 43-64.

Adams, F., & Aizawa, K. (2007) *The Bounds of Cognition*. Blackwell, Oxford, UK.

Butler, K. (1998) *Internal Affairs: Making Room for Psychosemantic Internalism*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Clark, A. (1997) *Being There: Putting Brain, Body and World Together Again*. MIT Press, Cambridge, Massachusetts.

Clark, A. (2003) *Natural-Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*. Oxford University Press, Oxford, UK.

Clark, A. (2008) *Supersizing the Mind: Embodiment, Action, and Cognitive Extension* Oxford University Press, USA.

Clark, A., & Chalmers, D. (1998) The Extended Mind. *Analysis*, 58(1), 7-19.

Golbeck, J. (2008) Trust on the world wide web: a survey. *Foundation and Trends in Web Science*, 1(2), 131-197.

Hollan, J., Hutchins, E., & Kirsh, D. (2000) Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2), 174-196.

Hutchins, E. (1995) *Cognition in the Wild*. MIT Press, Cambridge, Massachusetts.

Jorgensen, C., & Binsted, K. (2005) *Web Browser Control Using EMG Based Sub Vocal Speech Recognition*. 38th Annual Hawaii International Conference on System Sciences (HICSS'05), Big Island, Hawaii.

Kirsh, D. (1995) The intelligent use of space. *Artificial Intelligence*, 73(1-2), 31-68.

Kirsh, D., & Maglio, P. (1994) On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18, 513-549.

Kuhn, G., Amlani, A. A., & Rensink, R. A. (2008) Towards a science of magic. *Trends in Cognitive Science*, 12(9), 349-354.

Rowlands, M. (2003) *Externalism: Putting Mind and World Back Together Again*. Acumen Publishing Limited, Chesham, UK.

Rumelhart, D., Smolensky, P., McClelland, J., & Hinton, G. (1986) Schemata and sequential thought processes in PDP models. In D. Rumelhart & J. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 2* (pp. 7-58). MIT Press, Cambridge, Massachusetts.

Simons, D. J., & Levin, D. T. (1997) Change blindness. *Trends In Cognitive Sciences*, 1(7), 261-267.

Smart, P. R., Engelbrecht, P. C., Braines, D., Hendler, J. A., & Shadbolt, N. R. (2008) *The Extended Mind and Network-Enabled Cognition.* School of Electronics and Computer Science, University of Southampton, Southampton, UK. (Ref: ITA/P12/NEC)

Sterelny, K. (2004) Externalism, Epistemic Artefacts and the Extended Mind. In R. Schantz (Ed.), *The Externalist Challenge: New Studies on Cognition and Intentionality* (pp. 239-254). de Gruyter, Berlin, Germany.