We Have Always Been ... Cyborgs

Andy Clark, *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press, 2003. Pp. viii + 229. US\$26.00 HB.

This book is an exploration and popularisation of 'active externalism' or the 'extended mind' hypothesis (Clark, 1997, Clark and Chalmers, 1998, Dennett, 1996, Donald, 1991, Hutchins, 1995). I begin with some background about active externalism and the two principal arguments for holding it: the *parity argument* and the *complementarity argument*. This book develops the complementarity argument. I think that both arguments are problematic. But I go on to suggest that this fascinating book provides us with preliminary grounds for endorsing a version of active externalism that is stronger and stranger than the one Clark actually defends.

Active externalism is the belief that mind extends into the world, beyond the skin-and-skull boundary. Clark and Chalmers (1998) say that *cognitive processes* extend into the world when we use pen and paper to work something out, or when we use a computer, or even when we use language, which Clark thinks was the first technology. Also, *cognitive states* extend into the world when we use physical objects, or data-structures such as chips or CD-ROMs, as *external memory stores*.

In the past, Clark has given us two reasons for subscribing to this thesis. The first (and original) reason was the *parity argument*: if something counts as cognitive when it is performed in the head it should also count as cognitive when it is performed in the external world. Suppose you have to rotate images of geometrical shapes on a computer screen. You can rotate them using a neural implant in your head, or you can rotate them using a 'rotate' button in the world. Presumably we will say that the implant case is cognitive – so why isn't the button case cognitive as well? Clark and Chalmers say that epistemic credit is due where epistemic actions are performed, regardless of whether they are performed in the head or in the world.

This, however, only covers *cognitive processes* and Clark and Chalmers admit that the processes might be in the world whilst all of our "truly mental states – experiences, beliefs, desires, emotions, and so on" might be in the head (1998, p. 12). They therefore take the parity argument a stage further and argue that *cognitive states* can be constituted partly by features of the environment.

This brings us to the strange case of Otto's notebook. Otto suffers from Alzheimer's disease. He hears that there is an exhibition at the Museum of Modern Art. He consults his notebook, which says that the Museum is on 53rd Street. He walks to 53rd Street and goes to the Museum. Clark and

Chalmers say that the notebook plays the same role for Otto that biological memory plays for the rest of us. It just happens that “this information lies beyond the skin” (p. 13).

Clark and Chalmers’ claim is not only that Otto’s belief is out there in the world. It is that he believed the Museum was on 53rd Street before he looked it up, courtesy of the “functional isomorphism” between the notebook entry and a corresponding ‘entry’ in biological memory. This parity argument is in effect a challenge: if we can come up with a non-question-begging, relevant difference between the entry in the notebook and the entry in biological memory we should say what it is.

Adams and Aizawa (2001) take up the challenge. They say that one thing that characterises a belief in the head, as opposed to a belief in a notebook, is original intentionality – intrinsic, non-derived content. Beliefs in notebooks derive their meanings from conventions or social practices, which ultimately derive from the capacities of cognitive agents. Beliefs in the head, on the other hand, are not derived from anything. They have intrinsic, non-derived content. If we take intrinsic content to be the mark of the mental (which they say is a “rather orthodox theory of the nature of the cognitive”, p. 52) then it just so happens that everything that is mental is in the head.

In his *Metascience* review of Clark’s 1997 book, *Being There*, Gerard O’Brien (1998) pointed to other relevant differences. Most obviously, beliefs in the head are causally active. My belief that there is an exhibition at the Museum of Modern Art, plays a causal role in my going to the Museum. Entries in notebooks, on the other hand, are causally passive: “The information they encode does not do any work unless we bring them under the gaze of our perceptual equipment. At this point the recorded information does become causally active, but only because it is now re-coded elsewhere – namely, inside our skulls” (O’Brien, 1998, p. 82).

In his reply to O’Brien, Clark admits that there is a “potential tension” between two components of the extended mind story. The parity argument stresses “the way that extra-neural elements can play a role similar to internal ones (as in talk of external memory, etc.)” (Clark, 1998, p. 99). But in *Being There* he had provided another, “more interesting and plausible argument” that turns on the way “external elements may play a role different from, but *complementary to*, the inner ones’ (ibid., my emphasis) This is the *complementarity argument*, and it now moves into the ascendancy. Clark says: “The argument for the extended mind *turns primarily* on the way disparate inner and outer components may co-operate so as to yield integrated larger systems capable of supporting various (often quite advanced) forms of adaptive success” (ibid., my emphasis). This is the argument of *Natural Born Cyborgs*, to which we now turn.

From Bioborgs to Cyborgs

Clark starts with Clynes and Kline's original 1960s vision of a cyborg. This was "a human agent with some additional, machine-controlled, layers of automatic (homeostatic) functioning, allowing her to survive in alien or inhospitable environments" (p. 32). Clark thinks that this vision was too narrow, because it "restricted the imagined cyborg innovations to those serving various kinds of bodily maintenance": the real human-machine symbiosis "expands and alters the shape of the psychological processes that make us who we are". It provides an "array of resources to which biological brains, as they learn and grow, will *dovetail* their own activities". This leads to the creation of "extended computational and mental organisations: reasoning and thinking systems distributed across brain, body, and world" (pp. 32–33).

This complementarity claim has two intimately interrelated components, which I shall call the 'bioborg' and the 'neural opportunism' arguments. The bioborg argument is that we are already bioborgs (my term), with modular brains and on-board robot-like devices (let's call them 'biobots') that we can launch to do our bidding. This leaves us well placed for mechanical augmentation, for a meshing between our onboard abilities and technologies out there in the world.

Our bioborg nature is beautifully illustrated by Aglioti's extension of the Titchener Circles Illusion (p. 101). This illusion makes us *consciously* see a circle as larger than it really is, but when the circle is replaced by a plastic disk and we reach for the disk our thumb and forefinger form the right-sized aperture to pick it up. A plausible explanation is that we have two incoming visual path-ways and the information processing associated with one of them is fooled by the illusion whereas the information processing associated with the other is not. The information that comes in through the ventral stream is routed through consciousness, reasoning and memory. The information processing associated with this pathway is fooled. The information that comes in through the dorsal stream is devoted to the fine control of ongoing action. It is not routed through consciousness and the brain is not fooled. The brain does not report this incoming information to consciousness. It just gets on with the job, launching the biobot 'pick up the disk' routine.

The neural opportunism argument is that our brains use enduring features of the world as external memory stores that we can consult as needs dictate. When we scan ('saccade') around a room our eyes return to the same place time and again. They do this because *that is where the information is*. We do not need rich inner models. We only need a broad idea of what's out there, plus the ability to zoom in and retrieve detailed information when we need to – on a need-to-know basis. This gives us the best of both worlds. It is less demanding on memory but enables us to command a rich database

– that is *out there in the world*. This opportunism has operated in consort with our bioborg brains to create cognitive technologies that complement our basic cognitive abilities.

Clark thinks that language was the first such technology. Language gives us a “cognitive short-cut” (p. 70). Research on chimpanzees shows that the use of plastic tokens enables them to reduce high-order abstract problems to lower-order problems that their brains can handle. Chimps were trained to associate a plastic token, such as a red triangle, with any pair of identical objects (such as two shoes) and to associate a differently shaped token with any pair of different objects (such as a beer can and a banana). The chimps could then solve the more complex problem of categorising pairs-of-pairs of objects in terms of *higher-order* sameness or difference. They could classify shoe-and-cup (different) as *the same higher order relationship* as beer-can-and-banana (also different). They could do this because both pairs would get the same kind of token. All they had to do was to compare the tokens.

Clark thinks that words work in the same way. They label complex concepts and enable us to ‘freeze’ our thoughts. This enables us to *think about* our thoughts and ideas. (Is it a good idea? What are my reasons for believing it?)

Language was the first phase of our cyborg existence. We needed to be smart in order to develop it, but it was our ability to turn thoughts and ideas into new stable objects that “started the real cognitive snowball rolling” (p. 83). We were then drawn upwards in a virtuous circle as one technology led to another. We have upgraded our mindware from speech to writing, through increasingly flexible forms of printing, and are about to engage in an intimate relationship with machines.

Language is what Clarks calls a ‘transparent technology’. We are so well integrated with it that it is almost invisible in use. We are surrounded by such technologies (pens, watches, telephones etc.) and are rapidly developing new ones. Mobile phone technology will be extended to give us cheap, tiny cameras that will beam information to family and friends whilst we are shopping. Houses will have wall-mounted weather displays. Implants in our bodies will communicate with one another about where we are and how we are. Augmented reality will overlay our experience of the world with personalised information, beamed to us by satellite. Lost on campus we will enter ‘library’, don an eyeglass, and see a green arrow pointing to the library.

Clark claims that we are so well integrated with these technologies that the problem-solving system is the biological system plus the technology. Suppose that somebody asks you if you know the time. You say that you do and *then* you look at your watch. You say that you know it because you know

that you can easily find it out. What does the knowing, according to Clark, is the extended system of you-and-your-watch. Again, suppose you have a chip on your shoulder that gives you easy access to information about women basketball players. Clark says that there is no relevant difference between accessing the information on the chip and accessing it in long-term memory. You *know* the information because you have easy access to it, and what does the knowing is you-and-the-chip.

We need to throw two more things into the mix. The first is *constructive learning*: the cortex is plastic and the environment can build neural substructures ‘on the hoof’, altering our internal architecture. The second is our long childhood, which gives our adaptable brains time to mesh with out-there technologies.

Cognitive States Versus Cognitive Content

What are we to make of all this? Let’s go back to the parity argument. O’Brien says that beliefs in the head are causally active whereas diary entries are not. I think that this is a symptom of a greater divide –between a cognitive state and the content of a cognitive state. In one sense, my belief that there is beer in the fridge is a cognitive state. It is in my head. It is the state of *believing* that there is beer in the fridge. Beliefs in this sense can be strong and passionate, sincere or insincere, short-lived or long-lasting. They can be causally active. But in another sense, my belief that there is beer in the fridge is the *content* of my cognitive state. It is *what* I believe. It is a proposition that can be written down in a public, communicable symbolism that expresses not only the content of my belief but the content of yours as well. Beliefs in this sense can be true or false, tautologous or contradictory, subscribed to by one or many, but they cannot be strong and passionate, sincere or insincere, short-lived or long-lasting. Nor can they be causally active. My state of believing that there is beer in the fridge might cause me to go into the kitchen, but the *content* of my belief, the proposition, divorced from the cognitive state, cannot cause anything.

Another way of distinguishing between state and content is to show that different states can have the same content. We can believe and fear the same thing – that there is no more beer left in the fridge, for instance.

If we confuse these different senses of belief we will end up saying that a belief is sincere and tautologous, or that it is contradictory and four years old, which is like the old joke that Gilbert Ryle used as an example of a category mistake: “she came home in a flood of tears and a sedan chair”. Frege was one of the first to draw the state/content distinction: “a proposition may be thought, and again it may be true; never confuse these things” (Frege, 1967).

Clark and Chalmers begin by arguing that cognitive processes can extend into the world, but admit that “truly mental states – experiences, beliefs, desires, emotions, and so on” might be in the head (p. 12). Consequently, they take their argument a stage further and argue that cognitive *states* can extend into the world. But the distinction between state and content undermines their argument, because the natural thing to say is that cognitive states are in the head, even though their content may be in the world. Otto’s diary contains the *content* of his cognitive state, but the state itself (once he has read his diary) is in his head. Similarly with the complementarity argument: Clark says that cognitive states can mesh with cognitive technology, which can determine the content of the states. OK. But why does this mean that our *states* extend into the world?

Clark says that when we mesh with our technologies we form an extended system, and that this system is the new self. There is a weak and a strong claim here. The weak claim is that we form an extended system. The strong claim is that this extended system is the new self. Consider the strong claim first. One of Clark’s examples is piloting a commercial jet-liner, which he says is a task in which “human brains and bodies act as elements in a larger, fluidly integrated, biotechnological problem-solving matrix” (p. 25). But when we see Bert flying his Boeing overhead we do not say: “There goes Bert, with his four engines”. We do not think of Bert, when he is flying his Boeing, as Bert-and-his-Boeing (and just plain Bert, when he is not).

The weak claim is equally problematic. If I dig a hole in my garden with a spade, it is not the coupled system of me-and-the-spade that does the digging. My-spade-and-I do not get the prize for ‘best hole in the garden’. I get the prize, even though I could not have done the digging without the spade. By the same token, when I use pen and paper to complement my on-board pattern-recognition abilities, it is not the coupled system of me-and-the-pen-and-paper that performs the operation. It is just *me*, using the pen and paper. As Adams and Aizawa say, “mere causal coupling of some process with a broader environment does not, in general, thereby extend that process into the broader environment” (ibid., p. 56).

Now, it seems to me that we can abandon both of these claims and leave a version of the complementarity argument intact. We can still say that our on-board pattern-recognition abilities mesh with cognitive technologies (just as I mesh with my spade). We can even say that these technologies determine the content of our cognitive states. Reading an entry in my diary will determine the content of at least one of my cognitive states at the time. A central thesis that Clark wishes to defend is that we mesh with cognitive technologies, which determine the content of our cognitive states – as in the case of the diary. In this sense, of entwining and enmeshing, we can still be cyborgs, whilst still being old-fashioned skin-bags with our cognitive states inside our skulls.

Send Forth the Avatars!

Things are not so simple, however. I introduced the distinction between state and content to force us to ask whether cognitive *states* can extend beyond the skin-and-skull barrier. Although the lines of thought I have examined so far give us no reason to think so, Clark nevertheless provides other unsettling evidence that they can.

Let me explain my strategy here. This book provides us with good preliminary grounds for defending a version of active externalism that is stronger than the version Clark actually defends. That is, there is evidence in the book for entertaining such a position, even though Clark does not do so himself. Consequently I find myself in the strange position of criticising the claims he does make, whilst endorsing a stronger claim that he does not make. I may, of course, have misunderstood his reasons for advancing the evidence in question. My guess, however, is that he does not adequately distinguish between state and content and, consequently, does not ask whether cognitive states (not content, but states) can cross the skin-and-skull barrier. I am inclined to think that they can. Let us look at the evidence.

First, Clark relates Daniel Dennett's story, "Where Am I?". Dennett's brain is removed and put in a vat, but it is kept in radio contact with his body and senses. When the nurse wheels him in and he sees his brain in the vat he feels that he is *outside* the vat, looking at his brain. (His body is later trapped by a rockslide, entombed far beneath the Earth's surface. At first he feels trapped, but as the radio links give out he feels himself relocated into his brain.)

Clark also tells us about the two-feet-long nose. You are blind-folded and sit behind someone. A volunteer takes your right hand and uses it to tap and stroke the nose of the person in front of you, whilst tapping and stroking your nose with your left hand. If this is done in perfect synchrony there is a 50% probability that you will feel that your nose is two feet long. A similar scenario is that you look at an unattached, dummy hand. Somebody taps the dummy hand and your hand in perfect synchrony. Many subjects experience sensation in the dummy hand.

These experiments illustrate Ramachandran's Principle, which is that the brain depends on *perceived correlations* to construct a model of (and hence a sense of) bodily bounds and locations. In the nose case it depends on the correlation between tapping a nose and feeling your nose being tapped. Another way of saying this is to say that the brain will give us a sense of presence if it gets the right kind of feedback from its actions – if the feedback is fast and fluent, if actions are reciprocated, if there is flow and continuity. If these conditions are satisfied our brains can project feeling and sensation beyond the limits of the biological body, to get a new kind of body image that includes

non-biological components. Clark says that the big lesson is that “embodiment is essential but negotiable” (p. 114). Humans are never disembodied intelligences. All the evidence underscores the importance of touch, motion and interaction.

It seems to me that embodiment is *highly* negotiable. Suppose that Dennett’s eyes are divorced from his body but kept in radio contact with his brain. His brain is in the vat and his body is in bed. The nurse wheels his eyes in on a very small trolley. He now sees his brain in the vat and his body in bed. His cognitive states and his sense of himself are outside his body, where his eyes are.

Similarly, when you experience a sensation in the unattached dummy hand you experience a sensation that is outside your body. When you have a pain in your foot the pain really is in your foot. Nobody wants to say that it is in your brain. By the same token, when you feel a sensation in the dummy hand the sensation really is in the hand, not in your brain.

Fast, fluent feedback, reciprocal action, and flow and continuity are constraints that, in principle, are easily satisfied. Suppose we send out sensory avatars. Our eyes go to Titan and our ears go to Europa, whilst our brains stay at home in their concrete bunkers. Our sense of our location now depends on the sensory modality we focus on. We will need feedback and reciprocity, but this can be achieved. In fact we only need the *semblance* of feedback and reciprocity. This can be achieved by emulator circuits, which Clark discusses (p. 106). We use such circuitry in our everyday interactions with the world, to speed up our responses to external stimuli by anticipating what they are going to be. The input from our remote sensors can be veridical even if the reciprocity is somewhat contrived. What matters is that we will have a perfectly veridical sense of ‘being there’, even though our physical presence is minimal.

I suspect that our physical presence can become vanishingly small, with multiple sensory avatars (located in different places) enabling us to switch our sense of location at will. And why have real avatars? Why not have virtual ones, which would give us the input that we would get if we (or they) were at the location? This information might be retrieved from memory or it might be a prediction based on a general knowledge of the area (think of the complex calculations involved in predicting the weather; then think of a similar science of *perception prediction*). This information, suitably packaged with feedback and reciprocity, would give us a feeling of being present at the location.

What about our bodies back home? Our sense of ourselves would go outwards whilst the information-processing wetware remains behind. From our distant eyries on Titan we might see relays of our ancient brains and think how quaint they are. But if we achieved this level of technology it is unlikely that we would trust our survival to anything as unreliable as our ancient

wetware. If the brain really is a tool-kit of devices we will probably replace them with more reliable ones, and plenty of backups. These need not be stored in a single, central location. Some might go forth with the avatars, for ease of communication, whilst others remain behind. Some might belong to a single, shared resource, so that our minds mingle together (think of the World Wide Web). We can imagine these things. Clark says that our sense of ourselves is a construct. I do not see why our sense of ourselves, or the location of our cognitive states, has to be tied to our bodies. Or that we will need bodies in any strong sense, since our components can be distributed over many locations.

These are heady issues (or not . . .). They arise when we ask what I think is the key question: can cognitive *states* be outside the head? Talking about cognitive content is not going to cut the mustard.

I began this review by saying that *Natural-Born Cyborgs* is an exploration and *popularisation* of active externalism. This is a very readable book, pitched at a popular level. It is anecdotal and upbeat and taps into the spirit of the times. It is also a very funny book, in which Andy Clark's inimitable personality and endless optimism shine through. I am sure other readers will enjoy it as much as I did.

Terry Dartnall
School of Computing and Information Technology
Griffith University, Nathan, Queensland 4111, Australia