The Affirmation of Transhumanism: Uploaded Embodiment in a Virtual Environment

C.A.P. Linssen
Donders Institute, Nijmegen, The Netherlands
<charl@turingbirds.com>

17 October 2011

Abstract

Fear of death is one of our most primal emotions, with which we try to cope using various strategies. These serve to suppress, or 'deny' the existence of death.

Starting from the premiss that whole-brain emulation will advance to such a level that we can speak of uploaded minds, we explore the implications of an immortal existence in a virtual (computer) environment on our denial of death; in particular whether our fears can be mitigated by this new mode of existence.

1 Technologies for Human Enhancement

Briefly speaking, *transhumanism* (also *posthumanism*) is the improvement of the human condition using technology. This differs from using technology to cure the ill, in that the purpose of transhumanism is to 'upgrade' perfectly healthy individuals.

Improvements can be made in various areas. Genetic engineering can endow us with a longer lifespan, and a healthier life until our eventual demise. Nootropic drugs enhance cognition, whereas neural implants provide us with entirely novel abilities. Already, there are cortical implants which allow the blind to see and the deaf to hear. There is an equal interest (especially military) in 'reading the mind', for instance to let fighter pilots gain the upper edge over their opponents by bypassing kludgy joysticks and buttons for control [8]. Although primitive, these real-world examples of brain-computer interfaces give us a hint of the possibilities to come. This essay is primarily concerned with those transhumanist technologies that directly relate to the brain and its experience and perception of the world.

Judging which technologies qualify for the 'transhumanist' nomer can be difficult and arbitrary. If there is one key ingredient, then most certainly it is *integration*. This is the hallmark of a good, efficient tool. Use should be

Figure 1: The commercial QB telepresence robot by Anybots, Inc. (Mountain View, CA, USA)

unconscious; if it requires too much conscious attention, then it is not perceived as being a true extension of oneself—instead, we find ourself thinking more about the tool than the problem at hand. The classic example is of a physical instrument, such as a cane or a pencil: after having used it for sufficient time, the brain of the user incorporates the instrument into the personal egocentric space and its use becomes transparent.

A similar adaptation of the sense of self occurs in telepresence. Consider a camera mounted on a robot (Figure 1). If the camera image is projected onto our retinas, and at the same time we are in control of the gaze or movement of the camera (by moving the robot), we very quickly move our egocentric body space to that of the robot. This can be demonstrated by threatening the robot (for instance with a hammer) and measuring the galavanic skin response or brain activation patterns of the controlling subject, which are equivalent to when the person's real body is threatened [13].

2 Understanding the Brain

The brain has long been a source of inspiration for (applied) cognitive modelling, such as in the fields of Artificial Intelligence and Cybernetics. Especially the latter has always tried to make an accurate model of brain operation, and has in recent decades been spurred by developments in scanning technology, such

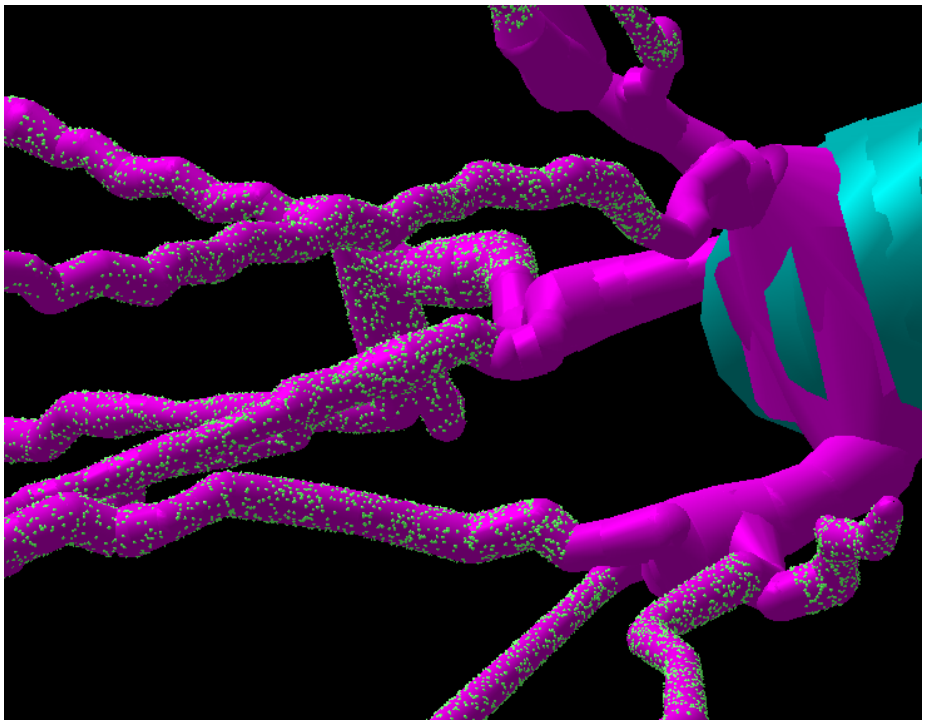

Figure 2: Screenshot of the Parallel Stochastic Ion Channel Simulator (PSICS) neuronal modelling tool [4], illustrating the level of detail in simulation presently achievable. Dots indicate individual ion channels on the dendritic tree.

as (functional) MRI and electron microscopy.

A converging interest in brain function comes from biology in general, and microbiology in particular. In this area, we see spectacular developments in imaging the neuron at a submicrometer resolution, allowing the *ultrastructure*—the microscopic details of synapses, branching dendrites, and so forth—of the neural circuit to be accurately mapped (Figure 2).

Modern neuroscience tells us that our brain is subdivided into many distinct faculties, such as those for motor control, vision, memory, and so forth [5]. Already, some of these are targeted for artificial prosthetics. This differs from the cortical implants mentioned in Section 1, which for instance bypass a failing eye or optic nerve and feed optical information directly into occipital cortex. Instead, neural prosthetics replace part of the brain itself. A prominent example is the hippocampus, the highest (most abstract) association cortex, and gateway of all episodic and semantic memories [6]. This part of the brain is to be replaced by a synthetic, silicon counterpart; that is, emulated using conventional computer hardware [8][2].

Neural prosthetics, which emulate in real-time the input-output behaviour and internal state of their biological counterparts, require not only knowledge

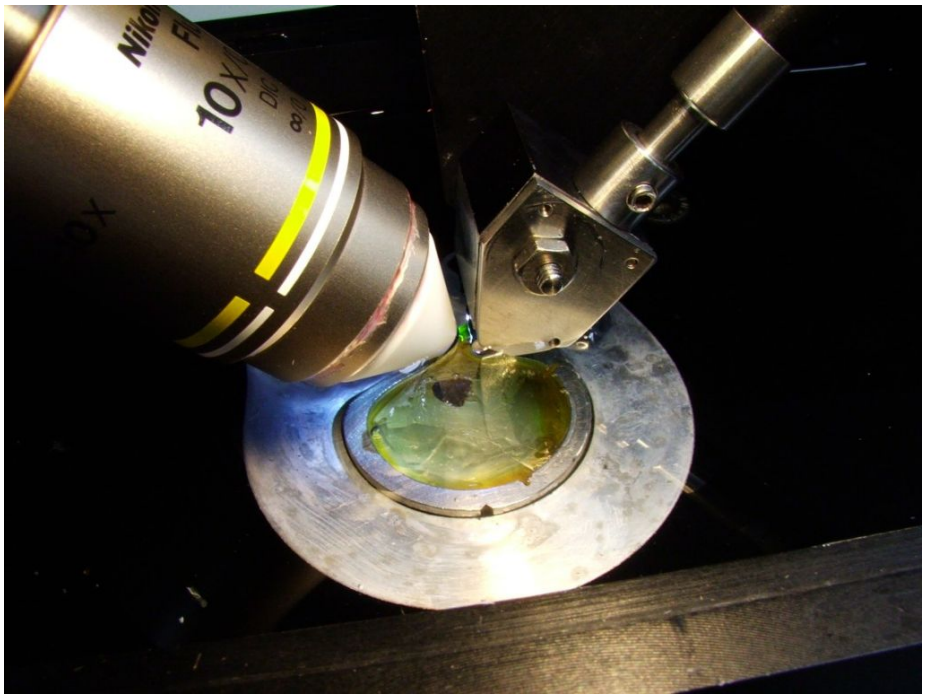

Figure 3: Knife-edge scanning microscope, Texas A&M University.

about the structural connectivity between cells and brain areas, but more importantly the *neural code*: how information is represented by firing patterns of neurons. Thanks to recent developments in microscopic imaging and voltage-sensitive dyes, the activity patterns of large populations can be recorded at single-cell resolution. This data can then be used to emulate the same circuits using non-biological hardware [10].

3 Whole-Brain Emulation

It stands to reason that with the continuously improving techniques for mapping physical neural connectivity, as well as highly accurate cell models (including e.g. gene regulatory networks and protein-protein interactions), it becomes possible to simulate (or emulate) the brain to a very high degree of accuracy. If the 'fidelity' of the simulation is high enough, this *whole brain emulation* is postulated to be an *in silico* incarnation of the original person. The process of transferring the mind from it biological, human brain substrate to the substrate of a computer is referred to as *uploading*.

At the time of writing, equipment for highly accurate virtual reconstruction of brain matter already exists. Figure 3 shows the Knife-Edge Scanning Microscope, developed at Texas A&M University [11]. A similar system, based on a

different operating principle but designed for the same purpose, is the Automatic Tape-Collecting Lathe Ultramicrotome (ATLUM) from Harvard University [7]. These systems benefit from the continuous improvements in microscopy, which result in increased reconstruction accuracy.

It is not yet known exactly what level of accuracy is sufficient to support whole-brain emulation: can we model neurons as abstract entities, described by a few mathematical equations, or do we need to model chemical concentrations for each compartment of the virtual neuron? [14]

The uploading postulate relies on the current neuroscientific consensus that all behaviour and experience is fully describable by physical laws. The outcome of these laws, or the evolution of the variables they describe over time, is computable. This holds true for all levels of emulation that could potentially be required[1]. The *Church-Turing thesis* can be summarized as follows: everything that is computable can be computed by a Turing machine. This implies that a computer supporting the emulation, with a particular architecture, can just as well be replaced by one with an entirely different architecture: perhaps one that does not use silicon, but carbon nanotubes as its computation elements. This implies that the uploaded mind has become *substrate independent*.

4 The Changing Meaning of 'Incarnation'

A considerable part of our brain is dedicated to interacting with the world: our *sensorium* and *motorics*. In the uploading process as described in Section 3, these features are faithfully copied into the digital model, perhaps up to the very point where the nerve fibres exit the spinal cord at the dorsal and ventral roots. We are clearly not only heavily invested in, but also intimately integrated with our basic sensorimotor processes. Thus, our perception is constrained by these structures, and we will need to live in an environment that matches them. We cannot live a fully disembodied existence: our senses have to be hooked up to *something* [5].

A *substrate* is the base upon with an organism lives. This definition commonly refers to the biological environment in which the organism can thrive. The substrate is directly perceived by us as we live in it: our brain is sitting right between our ears. Now consider what happens when a mind is transferred from the human brain substrate to the substrate of a computer. Although the computer chips that are supporting the whole-brain emulation are still sitting (tangibly) in a room somewhere, they are far removed from the direct experience of that person: a person's experience is not confined to that computer room. Thus, in an uploaded existence, the *perceived substrate* is different from our *actual* mind-substrate, which is silicon. The fact that we are *embodied* does not change: we still feel as though we are incarnate in a physically real body, even though it may not exist in the real world.

[1] At present, there is no evidence for quantum mechnical effects that play a functional role in the brain. See [14] for a discussion.

Figure 4: Two views of Second Life, showing the user's avatar and the virtual environment. Courtesy of Second Life user *RazWelles*.

The perceived substrate provides our direct experience of the world. This can, in general, be supplied either by a simulation or somehow be linked to the real world.

By the time uploading is feasible, we will have advanced robot implementations of a human body that can serve as the ultimate telepresence device for us [3]. It would allow humans to live in an environment where traditionally they could not, such as a planet without an atmosphere. Even though our mind simulation is taking place remotely, somewhere safe, our experience is instantiated in these robot 'avatars', which are free to roam within the range of the wireless communications link. Avatar robots extend the range of human action and allow us to escape the limitations of any biological body [12].

The alternative to instantiation in a real-world avatar, is to opt for instantiation in a virtual environment. Like the brain itself, this environment is simulated *in silico*. A familiar contemporary example is *Second Life*, a popular internet-based virtual world. Contrary to earthbound robotic avatars, these simulated avatars are not bound by the laws of physics. In the world of Second Life, one can for instance fly or instantly teleport from one location to another (Figure 4).

Because embodiment of a simulated brain is a matter of routing sensorimotor signals, these may at any point in time be re-routed to an avatar with different characteristics. The user could thereby select any desired body on demand. This is *multiple embodiment* [5][12]. In Section 1, we reviewed the remarkable capacity of the brain to adapt to new input, or to incorporate new faculties for the control of increasingly capable bodies. It is thus plausible that our brain would have no difficulty recalibrating itself to a novel avatar, be it a virtual embodiment, or a telepresence robot. As an example, consider your virtual body taking on the form of a bird, with the full sensory gamut of feathers and

wings, and the agility stemming from that, which cannot possibly be replicated so long as the user is in a human-like avatar.

5 The Denial of Death

Man is a 'hybrid species' [12]. Unlike any other, he has the capability for symbolic, abstract, intellectual experience. This contrasts with the physical, bodily realm that is apparently an inheritence from our evolutionary ancestors. One of our most elementary motives, that of self-perpetuation, stems from the reproductive purpose of a body in evolution. Consequently, this is one of our most fundamental and powerful drives.

The body places man in a more or less standardized species form. The intellect, on the other hand, is personal and achieved. Although our bodily makeup and basic biology are fixed, we have always tried to place them in a position of control, to try to exert intellectual control over them. However, in spite of our ability to greatly extend our capabilities using external tools, we are still bound by our finite bodies. These kludgy external tools rarely integrate well with our cognitive apparatus, limiting the freedom of our inner symbolic self, which strives for self-unfolding. Yet the inner self finds itself in an almost arbitrary body envelope, with its sexual, impermanent and 'digusting' biological nature (in reference to e.g. excrement).

Our intellectual life places an extra burden on us: now we do not only need to procreate in the physical world, but also to perpetuate ourselves in the symbolic realm; to pass on something of one's inner symbolic self. This is the terror of death: *'to have emerged from nothing, to have a name, consciousness of self, deep inner feelings, an excruciating inner yearning for life and self-expression—and with all this, yet to die.'* [1] Because of the dread and anxiety that this knowledge brings with it (only man has a real foresight of his impending demise), we develop strategies for repressing this knowledge, to prevent it from interfering with our daily functioning, thereby 'denying' death. However, these repressions are always imperfect and leave us vulnerable to physchological dysfunction and may in fact disrupt everyday life (e.g. in compulsions) [1].

6 Transhumanism and the Denial of Death

In 1973, Becker wrote: *'[...] how can an ego-controlled animal change his structure [...]? There is simply no way to transcend the limits of the human condition [...]'* Given what we have read in sections 3 and 4, it should be clear that this is no longer necessarily true. Technology *will* offer us a means to transcend the limits with which we have lived for some 200,000 years. What are the some of the implications of this historical transformation?

The fact that our bodies are mere avatars (be it in a simulated environment or

in the real world) implies that we could, as far as we can currently tell[2], be able to live indefinitely. In a simulated environment there would have to be no risk of dying in a car accident or plane crash. Even if we are embodied in an avatar that exists in the real, physical world, destruction of the avatar does not mean that the person dies, because the avatar can be controlled wirelessly from a safe location. Telepresence robots are employed in order to free their human users from the lethal repercussions of fate or bad judgement.

In the case of a mind upload, there is a dichotomy between the actual substrate (the silicon microchips) and the perceived substrate (an earthlike environment). A major question is thus what place the actual substrate will have in the mind of the person. On the one hand, the person is not as stuck to it as he is to his biological brain: one computer chip is as good as another (substrate independence). On the other hand, the person would still know about his inevitable dependence upon a physical structure in the world, with risk of destruction due to man, natural or cosmic disasters.

Reducing the probabilty of death is not a solution to the fear of death: *'The smallest virus or the stupidest accident would deprive a man not of 90 years but of 900—and would be then 10 times more absurd. [...] If something is 10 times more absurd it is 10 times more threatening.'* [1] However, this dependence could be mitigated by an advanced, multiple redundant setup: even a catastophe to the data centre would not wipe out the person. It therefore seems likely that the ultimate reliance on physical structures would not be given much significance in the face of the simulated environment; at least not nearly as much significance as our reliance on our present physical substrate.

In addition to not having to rely as much on physical structure for being, we have also seen in Section 4 that we can exercise a great amount of control over both our simulated and real-world avatars. A biological body is experienced as constraining because of its limitations: it is fixed and cannot be molded according to a person's wishes. Contrast this with either an advanced real-world robot avatar or a virtual body, especially the latter of which the form is completely subservient to the demands of the symbolic self. A virtual body has no need for 'disgusting' biological features such as defecation and procreation. The symbolic self gets to have full, ultimate control over its embodiment.

Existence in a silicon substrate would once and for all shift the emphasis from gene survival to pattern survival: *'The pattern is what's important, not the substrate.'* [15] Our relation to our bodies would radically change. Although it is hard to predict the full effect of this transformation on our complex psychology, clearly it will seriously diminish, if not abolish, our fear and denial of death, and all the imperfect repression mechanisms it breeds.

[2]I.e. with our current neuroscientific knowledge; perhaps the architecture of our brain is somehow fundamentally unsuitable for a lifespan measured in millenia, with respect to e.g. self-interference of stored memories.

7 Beyond Transhumanism

Kierkegaard claimed that fear of death ('dread' in his words) can actually enrich life, rather than diminish it: *'[. . .] not that [faith] annihilates dread, but remaining ever young, it is continually developing itself out of the death throe of dread'*. [9] Without these forces at play, what will remain of our inner drives towards intellectual self-perpetuation? If our patterns remain indefinitely, we may no longer feel the need to create external projects to perpetuate our symbolic selves.

We are the sum of our parts. If we take these drives away, that part of our humanity will be abolished for good. What implications, if any, will this have for our 'humanness', the extent to which we are that what makes us human? How much is our character and culture shaped by our having to deal with death and our finite biological bodies?

The very fact that we exist indefinitely, in an artifical substrate, in virtual worlds—which have a certain arbitrary character—will surely lead to its own existential crises, and new ways to give meaning to existence. We can only know ourselves if we continue to explore and expand ourselves [12]. Human psychology has many dimensions: fear of abolishing our humanity on the basis of abolishing fear of death seems unwarranted.

We can conclude that transhumanism itself is the invented solution to the problem of death; transhumanism *is* the denial of death. Like so many belief systems before it, it is a mechanism that allows us to cope with our finitude. What sets it apart is that it may be the first belief system which can *actually* eliminate death, and the associated need for repression of death anxiety in the first place: transhumanism has the potential to fulfill its promises, at which point itself and other such belief systems are no longer necessary.

References

[1] Ernest Becker. The Denial of Death. 1973.

[2] Theodore W Berger, Robert E. Hampson, Dong Song, Anushka Goonawardena, Vasilis Z. Marmarelis, and Sam A Deadwyler. A cortical neural prosthesis for restoring and enhancing memory. *J. Neural Eng.*, 8, 2011.

[3] Rodney A. Brooks. Flesh and Machines: How Robots Will Change Us. 2002.

[4] Robert C. Cannon, Cian O'Donnell, and Matthew F. Nolan. Stochastic Ion Channel Gating in Dendritic Neurons: Morphology Dependence and Probabilistic Synaptic Activation of Dendritic Spikes. *PLoS Computational Biology*, 6(8), 8 2010.

[5] Andy Clark. Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence. 4 2003.

[6] Greenberg and Verfaellie. Interdependence of Episodic and Semantic Memory: Evidence from Neuropsychology. *Journal of the International Neuropsychological Society*, 16:748–753, 2010.

[7] K.J. Hayworth, N. Kasthuri, R. Schalek, and J.W. Lichtman. Automating the

Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microscopy and Microanalysis*, 12, 2006. URL `http://www.mcb.harvard.edu/lichtman/ATLUM/ATLUM_web.htm`.

[8] Hannah Hoag. Remote Control. *Nature*, 423, 6 2003.

[9] S. Kierkegaard. The Concept of Dread. 1944.

[10] Henry Markram. The Blue Brain Project. *Nat Rev Neurosci.*, 7:153–160, 2 2006.

[11] B.H. McCormick. Development of the Brain Tissue Scanner. 3 2002. URL `http://research.cs.tamu.edu/bnl/kesm.html`.

[12] Hans Moravec. Mind Children: The Future of Robot and Human Intelligence. 1 1990.

[13] V.S. Ramachandran and S. Blakeslee. Phantoms in the Brain: Probing the Mysteries of the Human Mind. 1998.

[14] Anders Sandberg and Nick Bostrom. Whole Brain Emulation: A Roadmap. 1998. URL `http://www.fhi.ox.ac.uk/__data/assets/pdf_file/0019/3853/2008-3.pdf`.

[15] Frank Tipler. The Physics of Immortality: Modern Cosmology, God and the Resurrection of the Dead. 1994.