METAPHILOSOPHY

Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK, and 350 Main Street, Malden, MA 02148, USA
METAPHILOSOPHY
Vol. 42, No. 5, October 2011
0026-1068

EMERGING TECHNOLOGIES AND THE FUTURE OF PHILOSOPHY

PHILIPPE VERDOUX

Abstract: This article examines how a class of emerging technologies—specifically, radical cognitive enhancements and artificial intelligence—has the potential to influence the future of philosophy. The article argues that progress in philosophy has been impeded, in part, by two specific constraints imposed on us by the natural architecture of our cognitive systems. Both of these constraints, though, could in principle be overcome by certain cognitive technologies currently being researched and/or developed. It surveys a number of these technologies, and then looks at a particular metaphilosophical stance (called "inflationism") that advocates amplifying the abilities of philosophers rather than reducing the ambitions of philosophy, given the apparent "teleological gap" between philosophy's ultimate goal (i.e., "the truth") and the limited capacities of our evolved mental machinery.

Keywords: artificial intelligence, cognitive closure, cognitive enhancement, emerging technologies, metaphilosophy, neurotechnologies.

1. Introduction

> Any useful idea about the future should appear to be ridiculous.
> —Jim Dator (1996)

No article on the future of philosophy—or, for that matter, any other field of intellectual inquiry (Verdoux 2010a, 2010h)—can afford to ignore the influence that emerging technologies are likely to have on the discipline. I refer here to a panoply of (what I broadly term) "cognitive technologies," some of which are only now becoming visible on the horizon of technological possibility; if realized, these technologies will bring about significant changes in the nature of scholarship. The idea is, in brief, that because philosophizing is an intellectual activity, any modification of our intellectual abilities will entail a corresponding change in our capacity to philosophize.[1]

[1] Bostrom makes a similar point about ethics: "To the extent that ethics is a cognitive pursuit, a superintelligence could do it better than human thinkers. This means that questions about ethics, in so far as they have correct answers that can be arrived at by reasoning

In this article, I suggest that there is good reason for thinking that the antecedent "if realized" will indeed be satisfied. Furthermore, I argue that once radical cognitive technologies are realized, the venerable field of philosophy will be transformed in (at least) two respects. Summarily put, progress in philosophy has been impeded by two distinct features of the human mind that impose *in principle* limitations on what we can and cannot understand.[2] On the one hand, while the territory of collective knowledge continues to expand at (something approximating) a geometric rate, the unenhanced human mind remains more or less fixed and finite; this yields the "problem of size." On the other hand, a number of canonical problems in philosophy may, according to transcendental naturalists like Colin McGinn, have solutions that are permanently beyond our epistemic reach; this yields the "problem of type" (see McGinn 1989, 364; Verdoux 2010a).

These two problems are not, of course, peculiar to philosophy; for example, advanced physics appears to be approaching, or have already reached, the absolute limits of human comprehension (Barrow 1999; see section 3.2 below). Nonetheless, both problems stand as principled sources of ignorance in philosophy, given that philosophy (a) is interdisciplinary, as a result of philosophical naturalism and the recent proliferation of "philosophies of *X*," and (b) tackles some of the most recondite problems, such as the mind-body problem, found in any area of human inquiry. The next logical step, then—one that no philosopher has yet taken—is to consider how the advanced cognitive technologies now being developed could possibly, and will probably, allow us to confront and overcome the two aforementioned problems. This article is thus an exploration of the possible and probable ways that such technologies will change our ability to *do* philosophy, and thereby shape philosophy's future.[3]

2. Cognitive Technologies

> When a distinguished but elderly scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong.
> —Arthur C. Clarke (2000)

The most significant changes to how scholarly work is done will come about through the creation of "posthumans" (for lack of a better term), or beings whose cognitive, emotional, or health-span-related capacities will, by

and weighting up of evidence, could be more accurately answered by a superintelligence than by humans" (Bostrom 2010, 280).

[2] See Verdoux 2010g for an extended critique of progressionism.

[3] I will attempt to remain silent about the normative issue of *preferability*—the third *P* in the "three *P*s" of future studies. In section 3, though, I discuss Walker's (2002) normative metaphilosophy, which suggests a future that is most preferable.

stipulation, "greatly [exceed] the maximum attainable by any current human being without recourse to new technological means" (Bostrom 2008). I would urge here that whatever one's initial reaction to the notion of posthumanity might be, the idea need not be the product of "irresponsible fantasizing"[4] but can be given an intellectually respectable interpretation.

Consider, on the one hand, the fact that cyborgs, or "biotechnological hybrids" (see Clark 2004), already populate the world today. They are the result of silicone breast implants, cochlear implants, pacemakers, psychoactive pharmaceuticals, deep-brain stimulation (DBS), myoelectrically controlled prosthetics, brain-machine interfaces (BMIs), powered exoskeletons, external information-processing devices, glasses, clothes, and so on (Bostrom and Sandberg 2009; Clark 2008; Ihde 1990; Moreno 2006; Lebedev and Nicolelis 2006).[5] It is, in fact, increasingly the case that human *adaptation* to our environment (the process whereby our phenotypic features come to adaptively complement the relevant set of environmental factors) is occurring through technological modification of both the niches in which we live (Odling-Smee, Laland, and Feldman 2003) *and* our own bodies/minds (Clark 2004, 2008; Verdoux 2010b). That is, we are organism-artifact hybrids evolving *in* and *to* an increasingly human-built environment through the intentional action of a literal designer (Verdoux 2010c, 2010d, 2010e).

On the other hand, it is helpful to recall that *Homo sapiens* (which our future-oriented ancestors might have called "post-australopithecines") does not constitute an evolutionary *telos*, or finalistic end point, at the apex of the phylogenetic tree. Although natural selection in the industrialized world has, as a result of modern medicine, become more or less causally impotent (Balter 2005; Curry 2006), there are still genetic changes occurring in the developing world (through selection) and as a result of subsidiary mechanisms like drift and sexual selection.[6] One geneticist, in fact, has suggested that a "plausible scenario" of the future involves the human species splitting into a "robust" and a "gracile" group, just as the australopithecines did (Curry 2006). The relevant point for present purposes is that, given both the trend of cyborgization and the ateleological nature of evolution, it is almost certain (i) that there *will* be a species of beings that succeeds modern humans[7] (the only remaining twig on the

[4] I borrow this locution from Hofstadter (2005, 181), who uses it to describe the "Kurzweil-Moravec scenario," i.e., that the technological singularity is impending and will entail radical changes to society. See also Chalmers 2010.

[5] Or, if one accepts the extended mind hypothesis, biotechnologies hybrids include coupled systems like a human and his or her Filofax, or iPad, and so on (see Clark and Chalmers 1998; Clark 2008).

[6] Not to mention the changes that *we* are introducing to our genomes through gene therapy.

[7] This does not necessarily mean that *Homo sapiens* will go extinct; we may coexist with our technological progeny.

hominid branch) and, furthermore, (ii) that these organisms will be at least in part constituted by nonbiological materials.

This being said, there are two distinct strategies for achieving superintelligence, both of which a growing number of mainstream philosophers are taking seriously (see Bostrom and Sandberg 2009; Schneider 2010; Chalmers 2010a). First, there are *cognitive enhancements* (of biological intelligence); and second, there is *artificial intelligence* (AI). We should also note here another important distinction, explicated further below, between "weak" and "strong" superintelligence: the former involves merely amplifying the capacities already possessed by the human mind, while the latter entails one or more qualitative changes (Bostrom 2003). In this section, I briefly consider enhancements and AI in turn; afterwards, in section 3, I expand on the two problems mentioned above and examine how the creation of posthumans, endowed with greater-than-human abilities for intellection, could transform philosophy's future.

2.1. Cognitive Enhancements

> Friends joke that I should get the iPhone implanted into my brain. But [. . .] all this would do is speed up the processing, and free up my hands. The iPhone is part of my mind already.
> —David Chalmers (2010b, ix)[8]

The term "cognitive enhancements" refers to a wide variety of technologies and practices; the common feature of all such phenomena is their efficacy in "[amplifying or extending the] core capacities of the mind through [the] improvement or augmentation of internal or external information processing systems" (Bostrom and Sandberg 2006, 311). (For critical comments on the normativity of the enhancement concept, see Verdoux 2010f.) Traditionally, bioethicists have contrasted the notion of *enhancement*, which targets the healthy individual, with that of *therapy*, which attempts to restore normal functions lost as a result of pathology, trauma, and so forth (Sandel 2004). Many philosophers and scientists, though, have come to reject a strong distinction between these concepts, arguing that in practice the dichotomy often collapses. As Bostrom and Sandberg write: "Cognitive enhancement of somebody whose natural memory is poor could leave that person with a memory that is still worse than that of another person who has retained a fairly good memory despite suffering from an identifiable pathology, such as early-stage Alzheimer's disease. A cognitively enhanced person, therefore, is not necessarily somebody with particularly high (let alone super-human) cognitive capacities" (2009, 312).

[8] Note that Chalmers's claim is only that the iPhone is part of his mind already *if* the extended mind hypothesis is correct.

One can further differentiate between "conventional" and "radical" enhancements. Conventional enhancements are, it turns out, rather ubiquitous throughout late modernity, although most are not recognized as such. Enhancements of this variety include such quotidian phenomena as education, an enriched environment, general health (especially, good "neural hygiene"), mental training, the Internet, and certain mind-augmenting drugs, to name a few.[9] Caffeine, for example, is a psychoactive stimulant known to improve alertness and memory (White 1998). According to one recent study, approximately 90 percent of adults in North America drink one or more cups of coffee each day (Lovett 2005); caffeine is thus a widespread and socially acceptable cognitive enhancer. Or consider the fact that an education often strives "not only to impart specific skills or information, but also to improve general mental faculties such as concentration, memory, and critical thinking" (Bostrom and Sandberg 2009, 312).[10] In fact, the process of becoming educated "often produce[s] more permanent neurological changes than do drugs" (Bostrom and Sandberg 2009, 314).

While conventional enhancements have had a profound influence on scholarship over the millennia (even if this influence is not always obvious to us), here I am interested more in radical cognitive enhancements. In contrast to the former, this group consists of mostly experimental technologies, such as genetic engineering, nootropic pharmaceuticals, BMIs, and tissue grafts, to name a few (Walker 2002, 2008; Naam 2005).

At present, much of the research aimed at actualizing these possibilities is funded by the Defense Advanced Research Projects Agency (DARPA), as the bioethicist Jonathan Moreno discusses in *Mind Wars* (2006). According to DARPA, the explicit goal of research in (as they call it) "augmented cognition" (AugCog) is to "extend, by an order-of-magnitude or more, the information management capacity of the 'human-computer' combination by developing and demonstrating enhancements to human cognitive ability in diverse and stressful operational environments" (qtd. Moreno 2006, 51). DARPA, for example, has already developed a prototype "cognitive-feedback helmet" that enables one to remotely monitor soldiers' stress levels during combat (Moreno 2006, 53–54), and a number of pharmaceuticals are now being researched in an effort to engineer "fearless" and "sleepless" soldiers (Moreno 2006, 114–20).

One such drug is modafinil, originally developed to treat narcolepsy. According to recent studies, modafinil has a number of cognition-enhancing effects in the healthy individual: for example, normal test sub-

[9] See Williams 1997 for detailed discussion of Environmentally Mediated Intellectual Decline (EMID). Many technologies discussed in the extended mind literature would probably fall under the category "conventional." Otto's notebook, for example, is about as conventional as one can get (see Clark and Chalmers 1998).

[10] Consider how much logic improves one's ability to philosophize.

jects who took modafinil had better working memories, especially when performing more difficult tasks and when their natural performance abilities were lower (Muller et al. 2004). Furthermore, "on a larger battery of tasks, modafinil was found to increase forward and backward digit span, visual pattern recognition memory, spatial planning, and reaction time/latency on different working memory tasks" (Bostrom and Sandberg 2009, 317). Obviously, improvements in one's ability to perform contrived tasks in the laboratory do not always translate into practical skills applicable in the real world; modafinil might not help the philosopher construct a new theory of free will, for example, although it *has* become a popular enhancement drug on college campuses.

A more speculative possibility involves modifying the genes responsible for the structure of the human brain. For example, we know that the NR2B gene codes for NMDA receptors are integral in long-term potentiation (the neural process that underlies learning and memory). In a recent experiment, for example, researchers overexpressed the NR2B gene in (as they dubbed them) "Doogie" mice. The resulting transgenic rodents, whose NMDA function was enhanced, far exceeded their wild-type littermates in various spatial learning activities (Tsien 2000; Tang et al. 1999). Furthermore, we know that specific homeobox genes, such as X-Otx2 in the frog, control the morphogenesis of different regions of the brain (Walker 2002). And recent research has found a variety of human genes, including MCPH1, ASPM, CDK5RAP2, and CENPJ (Tang 2006), that are responsible for the large size of the human brain (or, more specifically, for the high "encephalization quotient" of our species). Thus, another possibility for cognitive enhancement involves manipulating these genes to engineer individuals with brains even bigger than those of modern *Homo sapiens*.

Finally, it is worth pointing out that, because producing cognitive enhancements (like philosophy itself) is a cognitive activity, any improvements in our ability to *think* will feed back positively to further enhance our ability to produce enhancements. This has already happened with conventional enhancements, and in fact it (appears to have) occurred at the evolutionary dawn of our species. That is, the first humans, *Homo habilis*, were also (very likely) the first to manufacture stone tools, or *lithics*. One explanation of the "great encephalization" that our lineage underwent thus posits that these tools provided a significant evolutionary advantage; as a result, natural selection began to weed out those individuals who were less endowed with the cognitive capacities needed to fashion effective tools. This led to humans with increasingly large brains. The earliest tools of the Oldowan may have been, therefore, the first cognitive enhancement technologies that humans created (albeit ones whose enhancive effects worked by altering the selective environment).

Let us now turn to the second possible route to superintelligence, which discards the biological core that enhancements aim to improve.

2.2. Artificial Intelligence

> Thus the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control.
> —Irving J. Good (1966, 33)

An "artificial intelligence" (AI) is a nonbiological system that "acts intelligently" in a particular environment. That is, broadly characterized, "what [the agent] does is appropriate for its circumstances and its goal, it is flexible to changing environments and changing goals, it learns from experience, and it makes appropriate choices given perceptual limitations and finite computation" (Poole, Mackworth, and Goebel 1998). Note that, according to Bostrom's definition of "posthuman" (given above at the beginning of section 2), posthumans are multiply realizable kinds: whatever its material constitution, all an object needs to instantiate the property "is a posthuman" is to satisfy the single condition specified above (Bostrom 2008). Thus, with respect to cognition, if an AI system's ability to "organize information" (see Bostrom and Sandberg 2009) greatly exceeds the maximum attainable by any current unenhanced human, then that system would *be* a (superintelligent) posthuman.

There are a number of possible approaches to the creation of artificially intelligent agents; some of these methods are "extendable," while others are not (although they may nevertheless *facilitate* the realization of "strongly" superintelligent posthumans). A quick list might include mind-uploading (also called "whole brain emulation"), artificial evolution (or "evolutionary robotics"), direct programming (as in GOFAI),[11] reverse-engineering the brain and then running the brain's algorithms on a computer, and, finally, machine learning. I will discuss a few of these in turn.

Consider, first, mind-uploading. This strategy involves replacing the neurobiological substance upon which the mind supervenes with a nonbiological, perhaps silicon-based or nanotube substrate (Kurzweil 2003). Such replacement may entail either destroying or preserving the original brain (which thus introduces a number of important metaphysical issues relating to personal identity; see Schneider 2010). One destructive method is the "microtome procedure." The idea here is to solidify the brain to make it sectionable; one could do this "either by perfusion with paraffin, or by freezing to liquid nitrogen temperatures" (Strout 1997). The brain is then sliced into sections sufficiently small to determine, through electron microscopy (or some other technique), the precise microstructure of the brain's component parts and their various interconnections. The resulting

[11] This acronym stands for "Good Old Fashioned Artificial Intelligence"; see Dreyfus 1992 for criticism.

data are then sent to a computer that simulates the sectioned brain, at which point "the patient finds herself or himself in a shiny new [artificial] body" (Strout 1997; see also Chalmers 2010a, 42, and Bostrom 2003).

Another method, variously termed the "nanoreplacement procedure" and "nanotransfer" (Chalmers 2010a), involves injecting nanoscale robots into the brain. These nanobots position themselves around an individual neuron, learn how to simulate the neuron's behavior, and then destroy and replace that neuron. The result is a functionally equivalent nonbiological component causally interacting with the network of biological neurons in which it is embedded. If this process is iterated enough times, the brain eventually becomes an artificial substrate whose causal properties are indistinguishable from those of the original organ. The subject's mind has thereby been uploaded.

It is worth mentioning that such computer simulations of brain circuitry could also preserve the original specimen. One could accomplish this through some not yet developed high-resolution scanning technology (but see Strout 2002), which would image the brain "with fine enough grain that neural and synaptic dynamics can be recorded" (Chalmers 2010a, 42). Although such technology is not yet available, Chalmers reminds us that brain imaging technologies are developing at an extraordinary rate; it is therefore not unreasonable to suppose that the necessary high-resolution scanning apparatuses will be available in the relatively near term.

Finally, it is also worth making explicit the philosophical thesis upon which the above possibilities are predicated, namely, that consciousness is an "organizational invariant." This means that the particular substrate implementing the mind is irrelevant *so long as* it possesses all the necessary organizational properties. If the entire nation of China, for example, decided to replicate with sufficient accuracy the causal-functional organization of a human brain for a single day, it would follow (at least on the interpretation of functionalism here adopted) that China would become the supervenience base of a genuine mind, given appropriate stimuli and output capabilities (see Block 1980).[12] Furthermore, this interpretation of functionalism is compatible with a variety of different philosophical theories about what consciousness *is*. As Chalmers points out, the thesis that consciousness *depends* on a certain causal-functional organization of the underlying physical system says nothing about what *constitutes* consciousness itself. Thus, "even a property dualist can in principle accept [the] functionalist theory construed [in this] way" (Chalmers 2010a, n. 24). At present, virtually all of contemporary cognitive science is founded on at least some version of functionalism—in particular, *computationalism* (see Lakoff and Johnson 1999).

[12] I simply cannot see how the notion that the population of China (or whatever) can become the supervenience base for mental states is any less fantastical than the notion that mental states supervene on the hundred billion or so cells linked together in the cranium. Both seem equally outrageous.

The above discussion has explored a number of different kinds of mind-uploading procedures; all of these involve *emulating* the micro-anatomy and physiology of the biological brain. But emulation is not an extendible procedure—that is, while an uploaded mind may be easier to *enhance* in various ways, and while an army of (possibly cloned) uploads may facilitate progress toward the subsequent invention of an extendable method, the process of brain emulation itself can only produce a mind as intelligent as the original (although the speed of cerebration may be much faster, depending on the hardware, resulting in superintelligence of the "weak" variety; see section 3.1 below). The realization of "strong" super-intelligence, in contrast, requires a method that can be iteratively applied, one that enables the AI systems it produces to *exceed* the human mind. In Chalmers's view—one that I share—the most promising extendible method is probably *artificial evolution*.

Artificial evolution is at the heart of evolutionary robotics, a research program that emerged in the 1990s and has since produced some truly amazing results. Before considering a specific example of an evolved robot, we should establish the experimental and theoretical framework in which much evolutionary robotics research is conducted. To begin, a population of "organisms" undergoing artificial evolution can consist of either virtual robots embedded in a physics-based simulated environment or actual robots situated in a real-world actual environment. (Or, in other cases, robots undergo simulated evolution and are then implemented in real-world robots for further observation.) On one approach, each robot consists of an artificial "genome," or sequence of binary characters that encodes a simple neural network (for example, eight sensory neurons connected to two motor neurons with no hidden nodes) whose synaptic connections are differentially weighted. While recent work has focused on enabling such networks to be ontogenetically plastic, most experiments to date have involved artificial genomes that map onto *fixed* networks of neurons (that is, they are only phylogenetically plastic).

Finally, individual robots are placed in their environments for specified increments of time; after time expires, their behavior is evaluated according to a "fitness function," such as "an individual gains fitness points by keeping as far away from predator robots as possible." The fittest genomes are then subject to random genetic mutations and recombination (between paired genomes) and copied into the next generation. Quite surprisingly, though, "just a few hundred generations of selection are sufficient to allow robots to evolve collision-free movement, homing, sophisticated predator versus prey strategies, coadaptation of brains and bodies, cooperation, and even altruism" (Floreano and Keller 2010, 1).

Consider the case of homing. In this experiment, which involved actual miniature mobile robots called "Khepera," researchers wanted to know whether "complex cognitive skills could be evolved by simply exposing robots to more challenging environments" (Floreano 2010). Specifically, the environment to which the author refers consisted of a square arena with one corner marked by a light above it and a black patch painted (in a quarter-circle shape) on the floor; this patch represented a "battery charger." The Khepera were then given twenty seconds of battery power unless they recharged their batteries by passing over the black patch. Finally, the fitness criterion stipulated that points would be assigned according to "the average rotational speed of the two wheels and distance from the walls" (Floreano and Keller 2010, 3); in other words, the fittest robots were those that moved the most and kept farthest away from the walls (relative to their "conspecifics"). It follows that robots that recharged their batteries could gain more total fitness points, since they would have more time to roam about.

Initially, the robots' movement was random, and individuals passed through the painted corner by accident. But, as Floreano reports, "after 240 generations, [. . .] we found a robot that was capable of moving around the area, [going] towards the charging station only 2 seconds before the battery was fully discharged, and then immediately returning in the open arena" (Floreano 2010). What is most extraordinary about this experiment, though, is not just the apparently intelligent homing behavior of the evolved Khepera but that its network of artificial neurons acquired a single unit whose activation depended on the individual's "position and orientation in the environment." That is, "this neuron encoded a spatial representation of the environment" (Floreano 2010), rather like the specialized "place cells" and "head-oriented cells" in the hippocampus of the rat (Floreano and Keller 2010).

Evolutionary robotics has accomplished much in its roughly twenty-year existence, and we can expect more from it in the near future. If minds *are* organizational invariants, and we gain the computing power in the future to run elaborate simulations of artificially evolving organisms, then it seems quite plausible that AI systems with intelligence greater than or equal to that of modern humans could be actualized.

3. Progress in Philosophy

> You do philosophy with the mind you have, not the mind you might want or wish to have at a later time.[13]

There are numerous identifiable impediments to progress in philosophy; some of these may be classified as institutional, financial, or communi-

[13] The allusion here is to former U.S. secretary of defense Donald Rumsfeld's statement that "you go to war with the army you have, not the army you might want or wish to have at a later time."

cative (the latter of which has been mitigated by inventions like the academic journal and the Internet; see Bostrom 2008). In the present section, I focus on two specific impediments that are *cognitive* in nature: the problems of "size" and "type." These problems impede progress because progress in philosophy (and most other scholarly fields of inquiry) is measured, generally speaking, in terms of *knowledge acquisition* (with "the truth" as the end goal), and the size and type conundrums each impose principled limitations on how much knowledge we as (i) individuals and (ii) a collective whole can possibly acquire. This suggests a distinction between our *relative* and *absolute* ignorance about the universe, as I discuss further below; both kinds of ignorance, and hence both problems, can in principle be overcome through the creation of advanced cognitive technologies. Let us thus examine these two problems in turn.

3.1. The Problem of Size

> In other words, complexity has the ability to "multiply" like a pair of rabbits in a meadow.
> —Theodore Modis (2002)

The human mind has intrinsic limits with respect to how much information it can retain in memory, how quickly it can process information, and so on. In the past, such limits were unimportant (or at least not as important) because the total amount of knowledge about the universe had by our ancestors—say, around the time of Plato—was less than the total amount of information a single individual could fit in his or her head. Our Hellenic ancestors simply didn't know about cells, proteins, atoms, quarks, evolution, genes, computers, plate tectonics, the dinosaurs, meteorites, iridium, Hadley cells, the solar system, and so forth. Epistemically speaking, the world in which they lived was much less complex (although, of course, many such civilizations had ornate mythologies to account for otherwise mysterious natural phenomena; but such mythologies do not qualify as knowledge).

But since the start of the scientific revolution (circa 1550, and ending around 1700), the piecemeal accumulation of knowledge about a wide range of phenomena, from the quantum mechanical to the chemical, biological, psychological, sociological, and cosmological, has increased exponentially—or something approximating it. The universe has turned out to be vastly complex, although possibly based on a finite number of simple nomological patterns, as the ongoing proliferation of specialties in both the sciences and humanities attests. Indeed, specialties—or "subfields" of study, where the prefix "sub-" can be recursively added as needed—are epistemically necessitated by our "finitary predicament"

(Cherniak 1990, 8).[14] That is, not only is there a temporal constraint on human intellectual activity[15] (which, as I discuss below, is cognitive in nature), our capacity to effectively retain information is constrained by the nature of our evolved mental machinery. Thus, as more is known about a given domain of phenomena, the disciplines dedicated to studying those phenomena must become proportionally narrow in scope; or, in different words, because of the fixed epistemic limits of the human mind, the researcher's focus must become increasingly parochial in order to keep the quantity of information contained within his or her (shrinking) discipline more or less the same.[16]

As Gregg Jacobs pointed out not too long ago (summarizing Woolfolk and Lehrer), "It was possible as recently as three hundred years ago for one highly learned individual to know everything worth knowing. By the 1940s, it was possible for an individual to know an entire field, such as psychology. Today the knowledge explosion makes it impossible for one person to master even a significant fraction of one small area of one discipline" (2003, 22). Or, more sententiously put: *everyone today knows almost nothing about most things* (see Verdoux 2010a).[17] That is to say, our relative ignorance, measured as the difference between what the individual knows and what the collective whole knows at a given point in time, is rapidly increasing,[18] as a result of (i) the limits intrinsic to the human mind, and (ii) the exponential expansion of collective human knowledge about the universe.[19] (In fact, this "ignorance explosion" has led some to worry

[14] This is not to deny the truism that *all* possible cognitive agents in our universe are in a finitary predicament (if only because of the universe's eventual "entropy death"); but it is to say that there may be crucial differences between the finitary predicaments of different kinds of sentient beings.

[15] A growing number of scientists are working to significantly compromise this limit, by developing powerful life-extension technologies. See de Grey et al. 2002.

[16] Or, as an unidentified person quoted by Ronald Wright puts it, specialists are "people who know more and more about less and less, until they know all about nothing" (Wright 2004, 29).

[17] This is precisely what makes John Hodgman's 2005 "almanac" so comical. The book's intentionally ponderous title is: *An Almanac of Complete World Knowledge Compiled with Instructive Annotation and Arranged in Useful Order by Myself, John Hodgman, a Professional Writer, in* ***The Areas of My Expertise****, Which Include: Matters Historical, Matters Literary, Matters Cryptozoological, Hobo Matters, Food, Drink & Cheese (a Kind of Food), Squirrels & Lobsters & Eels, Haircuts, Utopia, What Will Happen in the Future, and Most Other Subjects*.

[18] As the critic of technology Langdon Winner writes, "If ignorance is measured by the amount of available knowledge that an individual or collective 'knower' does not comprehend, one must admit that ignorance, that is *relative ignorance*, is growing" (Winner 1977, 283).

[19] Or, in Barrow's words: "As our knowledge deepens and widens, so it will take longer to reach a frontier. This situation can be combated only by increased specialization, so that a progressively smaller part of the frontier is aimed at, or by lengthening the period of training and apprenticeship" (Barrow 1999, 108). The suggestion I present in section 3 is *neither* of these, but involves increasing the cognitive resources of the individual so that he or

that, for instance, "the early creative period of a scientist's life will be passed by the time he or she has digested what is known and arrived at the research frontier"; Barrow 1999, 108).

These observations are relevant to our present topic because philosophy has become, especially in the Anglophone world, a thoroughly interdisciplinary field of study. There are several discernible reasons for this. One pertains to W. V. Quine's influence. For example, Quine famously challenged the inveterate distinction, championed by the logical positivists, between analytic and synthetic statements. (Analytic statements are true or false according to the meaning of the constituent words, whereas synthetic ones depend on contingent matters of empirical fact.) The result was a principled blurring of the boundary between science and philosophy: while science was previously thought to be concerned with the formulation of true synthetic statements about the world, philosophy was focused on clarifying concepts through the discipline-defining method of conceptual analysis—that is, the *a priori* activity of specifying the necessary and sufficient conditions that an entity X must satisfy to fall within a given concept's extension. After Quine, though, this tidy distinction between philosophy and science appeared to be untenable; philosophy and science are *continuous* domains of investigation.

In addition to challenging this "dogma of empiricism," Quine advocated "naturalizing" a specific region of philosophical investigation, namely, epistemology. The argument here pertains not to the analytic-synthetic distinction but to the dichotomy between description (which analytic and synthetic were both taken to be instances of) and normativity (see Searle 1999 for discussion). According to Quine, the aim of specifying a set of criteria according to which we can distinguish epistemically "good" beliefs from those that are "bad" (and thus determine which beliefs we *ought* to accept) has proven otiose—"a lost cause" (see Kim 1988). Thus, Quine argues that epistemology should be absorbed by psychology, an empirical discipline that strives to merely describe, in causal-nomological terms, how humans actually form their beliefs, and how those beliefs change over time, without offering any normative evaluations of those beliefs (that is, without attempting to say which are, and which are not, epistemically justified). This radical proposal is Quine's "replacement thesis" (Kornblith 1994).

Many philosophers have found Quine's prescription—to make epistemology a "chapter of psychology"—unsatisfactory (Kim 1988); nonetheless, as a result of Quine's work, a more modest version of *naturalism* has gained many champions in contemporary philosophy. Peter Godfrey-Smith, for example, advocates a version that "requires that we begin our philosophical investigations from the standpoint provided by our best

she may be more of an epistemic generalist while spending *less* time training or in an apprenticeship.

current scientific picture of human beings and their place in the universe. [. . .] Science is a *resource* for settling philosophical questions, rather than a replacement for philosophy or the source of philosophy's agenda" (Godfrey-Smith 2003, 154). But beginning our philosophical projects from the standpoint of our best current science requires one to *know about* our best current science. Thus, naturalism requires that philosophers grasp both their own field *as well as* whichever subfields of science might be in any way germane to the philosophical project at hand. The naturalistic philosopher must be scientifically educated and informed.

But this is not the only reason philosophy is interdisciplinary; there is also the multiplicity of different subject matters about which philosophers philosophize. According to Frederick Ferre, philosophy proper consists of epistemology (the "theory of knowledge"), metaphysics (which asks, "What exists?"), axiology (which asks, "What has value?"), and methodology (which "[deals] with proper critical methods in general") (Ferre 1995, 7). The questions that each of these subfields brings up can then be posed within and about *any* area of extraphilosophical research. The result is a "philosophy of X," as in the philosophy of language, the philosophy of biology, the philosophy of religion, the philosophy of art, the philosophy of . . . mind, law, mathematics, education, science, politics, history, time, technology, and so on. Again, insofar as good philosophizing about X requires one to know about both philosophy and X, the philosopher of X is engaged in a thoroughly interdisciplinary kind of scholarship.

The point to which this discussion has been leading is this: with the explosive growth of collective human knowledge in both the sciences and the humanities, interdisciplinary work has become increasingly difficult to do well. As Jacobs observed above, it may be impossible, given certain human constraints, to become an expert on any one area of study today, such as both philosophy *and* X. As a consequence, then, "philosophy of X" scholarship suffers, and the norms of naturalism become increasingly more difficult to follow in practice. (Or, in Sellars's metaphilosophical view, our ability "to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term" becomes excessively arduous when the broadness of "things" and "hangs together" exceeds the intellectual capacities of the human [Sellars 1962, 37]).

At the risk of belaboring the point, another way of putting this goes as follows. Because the breadth and depth of each individual's total knowledge are, to an approximation, inversely correlated—that is, the more one knows about any single topic, the fewer topics one knows about; and the more topics one knows about the less one knows about any single topic (Verdoux 2010a)—an increase in knowledge about X means, generally speaking, less knowledge about Y, and vice versa. And this inverse relation directly follows from the constraints of time and memory: the day is only so long, and the mental space in which to retain information only so capacious.[20] Such is our finitary predicament.

Some scientists working within the field of "cognitive informatics" have, in fact, attempted to quantify the human capacity to remember. According to Wang, Liu, and Wang (2003), the memory capacity of the human brain is approximately 10^{8432} bits. And, with respect to the constraint of time (or the information-processing speed of the human brain), Ray Kurzweil (2003) writes: "The human brain has about 100 billion neurons, with about 1,000 connections from one neuron to another. These connections operate very slowly, on the order of 200 calculations per second, but 100 billion neurons times 1,000 connections creates 100 trillion-fold parallelism. Multiplying that by 200 calculations per second yields 20 million billion calculations per second, or, in computing terminology, 20 billion MIPS."[21] The point here is that the human brain *does* indeed have an upper limit in terms of its computing power, as well as of its capacity to store information in memory.[22]

It follows from these considerations that one way to overcome the epistemic problem of size (or complexity) is to develop effective cognitive technologies. Such artifacts could target, for instance, the constraints of time and memory by (a) increasing the maximum speed at which a mind can perform calculations (thus "lengthening the day," so to speak), and (b) increasing the total capacity of our cognitive systems to encode, store, and retrieve information. Other constraints arising from cognitive phenomena like attention, understanding, and even creativity could also be, to some degree, overcome (see Orca 2009). An uploaded mind running on powerful hardware, for example, would be able to process significantly more information per increment of time than current humans can. And an AI with a superhuman "memory system" might explore interdisciplinary connections that no human could possibly entertain, since the AI would be able to master not only one but *multiple* domains of scientific or humanistic knowledge.

In sum, advanced cognitive technologies seem to offer one way of overcoming the increasingly urgent size problem. If such technologies *are* realized in the future, then interdisciplinary fields like contemporary "analytic" philosophy, with its growing number of "philosophy of X" subfields, should expect to undergo a radical transformation, as superintelligent

[20] As Chomsky states, just as humans have the capacity to solve certain problems, so too do they "lack the capacity to solve other problems, which will either be far too difficult for them to handle within existing limitations of time, memory, and so on or will literally be beyond the scope of their intelligence in principle" (Chomsky 1988, 149).

[21] Kurzweil then adds that, given Moore's law, "we'll have 20 billion MIPS for $1,000 by the year 2020" (although this leaves open the software question) (Kurzweil 2003).

[22] Furthermore, Moravec estimates that "a computer with 100 million MIPS [one million machine instructions per second] should be equal to humans in intelligence" (Walker 2002).

agents (or, intermediately, variously modified humans whose computational capacities are augmented) come to master knowledge-domains that we unenhanced members of *Homo sapiens* are barely capable of glimpsing.

3.2. The Problem of Type

> The organ sitting in our heads has not the size and power to comprehend everything that exists.
> —Colin McGinn (1993, 36)

Even if a single posthuman could master every theory established by contemporary science and philosophy (whenever "contemporary" is exactly), a question remains concerning the capacity of that individual to comprehend, in principle, every true theory about the cosmos. In other words, overcoming the problem of size through advanced cognitive technologies would only eliminate our *relative* ignorance; it would have absolutely no bearing on our *absolute* ignorance.

Thus, the superintelligent philosopher of the future might have complete knowledge of both philosophy and physics, for example, but still not be able to understand deeply the nature of the eight additional spatial dimensions posited by M-theory.[23] As Bostrom puts it, "Our cognitive limitations may be confining us in a Platonic cave, where the best we can do is theorize about 'shadows', that is, representations that are sufficiently oversimplified and dumbed-down to fit inside a human brain" (Bostrom 2005). In contrast to the problem of size, then, epistemic quandaries of this variety result from principled constraints on the *conceptual space* to which a cognitive agent, such as *Homo sapiens*, has access, given the evolutionary peculiarities of its mental machinery.[24] By analogy, there may be true theories accessible to a class of qualitatively distinct minds (with respect to us humans) that are permanently beyond our own epistemic reach, just as there are true theories accessible to the human mind that are forever beyond the reach of crickets. This is the phenomenon of "cognitive closure," a central idea of *transcendental naturalism*.

[23] As Dawkins insightfully writes, "[I want to pursue the point] that the way we see the world, and the reason why we find some things intuitively easy to grasp and others hard, is that *our brains are themselves evolved organs*: on-board computers, evolved to help us survive in a world—I shall use the name Middle World—where the objects that matter to our survival were neither very large nor very small; a world where things either stood still or moved slowly compared with the speed of light; and where the very improbable could safely be treated as impossible. Our mental burka window is narrow because it didn't *need* to be any wider in order to assist our ancestors to survive" (Dawkins 2006, 367–68).

[24] Again, the problem of size only pertained to the complexity of the problem, and thus the human ability to retain and process information. The type problem thus concerns our *conceptual* limitations, the size problem our *computational* limitations.

According to Colin McGinn, a major proponent of transcendental naturalism,[25] "a type of mind *M* is cognitively closed with respect to a property *P* (or theory *T*) if and only if the concept-forming procedures at *M*'s disposal cannot extend to a grasp of *P* (or an understanding of *T*)" (McGinn 1989, 350). Using Chomsky's (1976, 1988) phraseology, we may call any puzzle that we can in principle answer a "problem," and any puzzle that is not merely unsolved but *unsolvable* a "mystery."[26] We may furthermore distinguish between partial mysteries that we can intelligibly ask about but cannot possibly answer, and absolute mysteries that we cannot even be perplexed by.[27] For example, McGinn argues that the problem of explaining how qualitative experience arises from our fatty neural tissue presents theoretical difficulties that we will never be able to overcome. That is, we can be—and *are*—puzzled by this phenomenon, but we cannot explain it. Why? Because the encyclopedia of concepts to which we humans have epistemic access does not include the concepts needed to grasp the relevant (naturalistic) theory. While few would controvert the existence of absolute mysteries (that is, the idea that we are forever ignorant of certain facts, as well as ignorant about our ignorance of those facts; this involves a kind of second-order ignorance), the notion that there are askable questions that are in principle unanswerable remains a point of contention among philosophers (see, e.g., Dennett 1991).

Finally, as alluded to above, cognitive closure may also be relevant in fields of advanced science like particle physics. As McGinn observes, "One of the areas in which the possibility of cognitive closure looks most real is theoretical physics—quantum theory and the origin of the universe being the standard examples. The more advanced a theory becomes the more likely it is to approach the limits of what we can know" (McGinn 1991, 88; see Verdoux 2010a). Another candidate is the field of AI; this is, in part, why the strategy of *artificially evolving* intelligent agents in the real world or computer simulations is so appealing: it is, as Douglas Hofstadter puts it, an "anti-intellectual" approach (2005, 180). (Indeed, what is more *anti-intellectual* than natural selection itself?)

[25] I understand transcendental naturalism to be the view that *there exist mysteries that the human mind cannot understand*. Whether there are partial mysteries in addition to absolute mysteries is a separate question—one need not think that consciousness, for example, poses a permanently insoluble puzzle to be a transcendental naturalist. See McGinn 1993.

[26] Furthermore, one could also distinguish between "illusions" and "issues" (see McGinn 1993).

[27] The former pushes against the idea that, as Wittgenstein put it, "in order to draw a limit to thinking we should have to be able to think both sides of this limit (we should therefore have to be able to think what cannot be thought)" (Wittgenstein 2010, 27). Or, in Walker's words: "Just as only we can appreciate exactly what it is that a child fails to know or understand, so too, it seems, only creatures who transcended our understanding should be able to detail our limitations. Perhaps a full philosophical account of our epistemic limitations is not something we are in a position to formulate or even appreciate" (Walker 2002).

Thus, the idea here is that advanced cognitive technologies could, potentially, minimize not only our relative ignorance but our absolute ignorance as well. That is to say, the actualization of "strongly" superintelligent posthumans through the creation of radical cognitive enhancements or advanced AIs could result in a *redefinition* of the boundary between problems and mysteries—just as Darwinian evolution redefined this boundary for *Homo sapiens* (with respect to our phylogenetic ancestors).[28] Such a "conceptual revolution"[29] is plausible because *cognitive closure is a mind-relative phenomenon*; thus, a new kind of mind, brought into existence through natural or artificial means, might classify a given puzzle that we take to be mysterious as an easily soluble problem.[30] "Conceivable creatures," McGinn notes, "might invert the classifications we make with these concepts, finding consciousness and free will easy to penetrate and explain scientifically, while being quite mystified by the movement of the planets or the nature of digestion" (McGinn 1994).[31]

The possibility of redefining the mystery-problem boundary is especially relevant to philosophy's future because philosophy is a repository of many of the most abstruse problems that humans have ever articu-

[28] Carol Rovane, in a response to Colin McGinn's transcendental naturalism, very briefly makes a similar suggestion. She writes: "But it doesn't follow that the only cognitive limitations that we can overcome are those which are imposed by our theories. For it doesn't follow that we cannot overcome limitations that are due to the very cognitive capacities that make theorizing possible. For one thing, our imagination can be directed at technological as well as theoretical innovation. And with the right technology we can overcome at least some of the limitations of our cognitive capacities. For example, the temporal and spatial limitations imposed by our capacities for perception, memory and calculation can to a significant extent be overcome with the aid of prosthetic devices such as telescopes, logs and computers. So some of the limitations that are due to the nature of our cognitive capacities clearly can be overcome" (Rovane 1994).

[29] At one point, McGinn argues that such a conceptual revolution "is not a revolution our intellects can effect" (1993, 35). In another place, though, McGinn states that "progress [in philosophy] would require us to overcome [our] architectural or constitutive limitations—which is not going to be possible without entirely reshaping the human mind" (1993, 150). Furthermore, he writes that "the truths of philosophy [might be apprehended] only by means of intellectual prostheses; or only after substantial cerebral enhancement" (1993, 129). Indeed, this is precisely what inflationism advocates—cerebral enhancement—and what I am here arguing appears to be *probable*, if not merely possible, in the near-term future.

[30] As Bostrom notes, artificial intellects may not have humanlike motives or psyches. For example, "The cognitive architecture of an artificial intellect may also be quite unlike that of humans. Artificial intellects may find it easy to guard against some kinds of human error and bias, while at the same time being at increased risk of other kinds of mistake that not even the most hapless human would make. Subjectively, the inner conscious life of an artificial intellect, if it has one, may also be quite different from ours" (Bostrom 2010, 279–80). See also Schneider 2010, 13–14.

[31] Or, as Chomsky puts it, "A Martian scientist, with a mind different from ours, might regard this problem [of free will] as trivial, and wonder why humans never seem to hit on the obvious way of solving it. This observer might also be amazed at the ability of every human child to acquire language, something that seems to him incomprehensible, requiring divine intervention" (1988, 152).

lated.[32] McGinn identifies (or claims to have identified) a number of these, including the nature and identity of the self, the foundations of meaning, the possibility of free will, and the availability of *a priori* and empirical knowledge, as well as the issue of consciousness and the mind-body problem (McGinn 1993). If these *are* partial mysteries, as defined above, then the creation of qualitatively new kinds of minds through advanced cognitive technologies could provide solutions that we cannot in principle understand, even if they were explained to us in exhaustive detail. Philosophy for these novel beings would be quite different from what it is for us. Either way, there is little doubt that absolute mysteries abound—that is, mysteries that we cannot even be puzzled by, just as the mouse cannot be puzzled by the nature and existence of free will. A much richer and more complete conception of the cosmos could be attained by such minds—a conception that is a little more complete in the absolute sense.

4. Cognitive Technologies and Metaphilosophy

> There are many questions—and among them those that are of the profoundest interest to our spiritual life—which, so far as we can see, must remain insoluble to the human intellect unless its powers become of a quite different order from what they are now.
> —Bertrand Russell (1936, 153)

According to Walker (2002), there are at least four metaphilosophical approaches to dealing with the apparent gap between the cognitive capacities of *Homo sapiens* and the end goal of philosophical inquiry, namely, "the truth." First, one could simply deny that such a "teleological gap" exists. This "denialist" position is, apparently, what Donald Davidson espouses (Walker 2002), as well as scientists like Richard Feynman (arguably) and Bently Glass (Barrow 1999, 80–81). Glass, for example, claims that "we are like the explorers of a great continent who have penetrated to its margins in most points of the compass and have mapped the major mountain chains and rivers. There are still innumerable details to fill in, but the endless horizons no longer exist" (qtd. Barrow 1999, 81).

Second, the philosopher could adopt a kind of "stoic resolve" and acquiesce to the ostensible fact that our "primitive intellectual capacities"[33] will forever preclude us from ascertaining certain correct theories about the cosmos. In Thomas Nagel's words, "The world is a strange place, and nothing but radical speculation gives us hope of coming up with

[32] That is, abstruse *for us*.

[33] From Nagel 1989, 10. In addition to Nagel, Russell and McGinn seem to fall within this category.

any candidates for the truth. That, of course, is not the same as coming up with the truth: if truth is our aim, we must be resigned to achieving it to a very limited extent, and without certainty. To redefine the aim so that its achievement is largely guaranteed, through various forms of reductionism, relativism, or historicism, is a form of cognitive wish-fulfillment. Philosophy cannot take refuge in reduced ambitions. It is after eternal and nonlocal truth, even though we know that is not what we are going to get" (1989, 10).[34]

Third, the philosopher could, contra Nagel, choose to "take refuge in reduced ambitions," thus advocating that the end goal of philosophy be modified or adjusted to better fit the limited abilities of the human philosopher; Walker (2002) labels this the "deflationary response." Exponents of deflationism include the logical positivists (see, e.g., McGinn 1993, 149), Hilary Putnam (1992), and William James (1995). James (quoted in Walker 2002) writes: "I firmly disbelieve, myself, that our human experience is the highest form of experience extant in the universe. I believe rather that we stand in much the same relation to the whole of the universe as our canine and feline pets do to the whole of human life. They inhabit our drawing rooms and libraries. They take part in scenes [of] whose significance they have no inkling. They are merely tangent to curves of history, the beginnings and ends and forms of which pass wholly beyond their ken. So we are tangent to the wider life of things" (James 1995, 299).

The obvious fourth option is what Walker calls *inflationism*: "The idea, in a slogan, is that it is not we who ought to abandon philosophy, but that philosophy ought to abandon us. Consider that as a mere point of logic, if there is a gap between the telos of philosophy and humanity then there are at least two means to close this gap: either philosophy can be scaled down into something more human, or philosophers can be scaled up into something more than human. The idea would be to create better philosophers, ones more naturally suited to realizing the ambitions of philosophy" (Walker 2002).

I suspect that this position has not been considered much by philosophers because its normative component has appeared (and *was*) infeasible. That is, if ought implies can, and if inflationism asserts that philosophers ought to be "scaled up into something more than human," then the inflationary stance asks us to do the impossible; it is therefore untenable. But the ongoing revolution in genetics, nanotechnology, and robotics (the "GNR revolution") is changing the feasibility of inflation-

[34] Nagel further writes that "there is a persistent temptation to turn philosophy into something less difficult and more shallow than it is. It is an extremely difficult subject, and no exception to the general rule that creative efforts are rarely successful. I do not feel equal to the problems treated in this book [*The View from Nowhere*]. They seem to me to require an order of intelligence wholly different from mine. Others who have tried to address the central questions of philosophy will recognize the feeling" (1989, 12).

ism's prescription. There is, as section 2 attempts to illustrate, a vast panoply of diverse technologies currently being developed that is certain to change the cognitive contours of philosophical scholarship (and all other fields of research, for that matter). Philosophers like McGinn are thus wrong to "boldly speculate that this *will* be the state of philosophy in a million years time, give or take some streamlining and clarification: the fundamental disputes [e.g., about the mind-body problem] will still rage, as irresoluble as ever, as they have these many centuries" (McGinn 1993, 152). McGinn here neglects the effects that advanced cognitive technologies are likely to have on how the philosopher of the future goes about explaining explananda that are, for us, either *mysterious* or simply too complex to grasp mentally. McGinn falls victim to a failure of futurological imagination.

Finally, we should note that some phenomena are easier to predict than others. The location of the moon (relative to the earth, or some other astronomical body) at exactly 9:00 A.M. on January 1, 2045, for example, can be determined with a high degree of accuracy. In contrast, attempting to predict how the Dow Jones Industrial Average will close two weeks from today is notoriously difficult. With respect to such refractory phenomena, the future studies scholar Jim Dator observes that "what is often popularly, or even professionally, considered to be 'the most likely future' is, in all probability, one of the least likely futures" (Dator 1996).

A good dose of circumspection is thus warranted in any futurological investigation, such as the present one about philosophy's future. Still, it is not unreasonable, I believe, to extrapolate from current trends in cognitive technologies to the conclusion that the future will contain beings whose cognitive abilities will far surpass our own abilities. The world today is already populated by biotechnological hybrids, and given Moore's law we should soon have the computing power necessary to run elaborate simulations in which virtual robots artificially evolve over millions of generational iterations.[35] Given these salient trends, then, a proper exploration of the future of philosophy ought to take seriously the possible, and probable, effects that advanced technologies will have on the discipline.

711 W. Carver St.
Durham, NC 27704
USA
philippe.verdoux@gmail.com

[35] As Bostrom writes, "Given the enormity of the consequences of superintelligence, it would make sense to give this prospect some serious consideration even if one thought that there were only a small probability of it happening any time soon" (Bostrom 2003).

Acknowledgments

I would like to thank the Institute for Ethics and Emerging Technologies and Whitney A. Trettien for supporting my work. Also, special thanks to an anonymous *Metaphilosophy* referee for excellent feedback and criticism.

References

Balter, M. 2005. "Are Humans Still Evolving?" *Science* 309, no. 5732: 234–37.

Barrow, J. 1999. *Impossibility: The Limits of Science and the Science of Limits*. Oxford: Oxford University Press.

Block, N. 1980. "Troubles with Functionalism." URL: http://w3.uniroma1.it/cordeschi/Articoli/block.htm.

Bostrom, N. 2003. "Ethical Issues in Advanced Artificial Intelligence." URL: http://www.nickbostrom.com/ethics/ai.html.

———. 2005. "Transhumanist Values." URL: http://www.transhumanism.org/index.php/WTA/more/transhumanist-values/.

———. 2008. "Why I Want to Be Posthuman When I Grow Up." In *Medical Enhancement and Posthumanity*, edited by Bert Gordijn and Ruth Chadwick, 107–37. New York: Springer.

———. 2010. "Ethical Issues in Advanced Artificial Intelligence." In *Science Fiction and Philosophy: From Time Travel to Superintelligence*, edited by Susan Schneider, 277–83. Oxford: Wiley-Blackwell.

Bostrom, N., and A. Sandberg. 2006. "Converging Cognitive Enhancements." *Annals of the New York Academy of Sciences* 1093: 201–27.

———. 2009. "Cognitive Enhancement: Methods, Ethics, Regulatory Challenges." *Science and Engineering Ethics* 15, no. 3: 311–41.

Chalmers, D. 2010a. "The Singularity: A Philosophical Analysis." *Journal of Consciousness Studies* 17, nos. 9–10: 7–65. URL: http://consc.net/papers/singularity.pdf.

———. 2010b. "Foreword." In *Supersizing the Mind*. Oxford: Oxford University Press.

Cherniak, C. 1990. *Minimal Rationality*. Cambridge, Mass.: MIT Press.

Chomsky, N. 1976. *Reflections on Language*. London: Fontana.

———. 1988. *Language and Problems of Knowledge*. Cambridge, Mass.: MIT Press.

Clark, A. 2004. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford: Oxford University Press.

———. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

Clark, A., and D. Chalmers. 1998. "The Extended Mind." *Analysis* 58, no. 1: 7–19.

Clarke, A. C. 2000. “Hazards of Prophecy: The Failure of Imagination.” *In Profiles of the Future: An Enquiry into the Limits of the Possible*. New York: Indigo Paperbacks.
Curry, O. 2006. “The Bravo Evolution Report.” URL: www.kingkraal.com/bravo.pdf.
Dator, J. 1996. “Foreword.” In *The Knowledge Base of Future Studies*. URL: http://www.futures.hawaii.edu/dator/futures/foreword.pdf.
Dawkins, R. 2006. *The God Delusion*. New York: Mariner Books.
de Grey, A., B. Ames, J. Andersen, A. Bartke, J. Campisi, C. Heward, R. McCarter, and G. Stock. 2002. “Time to Talk SENS: Critiquing the Immutability of Human Aging.” *Annals of the New York Academy of Sciences* 959: 452–62.
Dennett, D. 1991. Review of McGinn, *The Problem of Consciousness. Times Literary Supplement* (May 10).
Dreyfus, H. 1992. *What Computers Still Can’t Do: A Critique of Artificial Reason*. Cambridge, Mass.: MIT Press.
Ferre, F. 1995. *Philosophy of Technology*. Athens, Ga.: University of Georgia Press.
Floreano, D. 2010. “Evolutionary Robotics: A Gentle Introduction.” URL: http://lis.epfl.ch/index.html?content=resources/documentation/EvolutionaryRobotics/index.php.
Floreano, D., and L. Keller. 2010. “Evolution of Adaptive Behavior in Robots by Means of Darwinian Selection.” *PLoS Biol* 8, no. 1: 1–8.
Godfrey-Smith, P. 2003. *Theory and Reality: An Introduction to the Philosophy of Science*. Chicago: University of Chicago Press.
Good, I. J. 1966. “Speculations Concerning the First Ultraintelligent Machines.” In *Advances in Computers*, edited by F. L. Alt and M. Rubinoff, 6: 31–88. New York: Academic Press.
Hodgman, J. 2005. *The Areas of My Expertise*. Boston: Dutton Adult.
Hofstadter, D. 2005. “Moore’s Law, Artificial Evolution, and the Fate of Humanity.” In *Perspectives on Adaptation in Natural and Artificial Systems*. Oxford: Oxford University Press.
Ihde, D. 1990. *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
Jacobs, G. 2003. *The Ancestral Mind: Reclaim the Power*. London: Penguin.
James, W. 1995. *Pragmatism: A New Name for Old Ways of Thinking*. London: Dover.
Kim, J. 1988. “What Is ‘Naturalized Epistemology’?” *Philosophical Perspectives* 2: 381–405.
Kornblith, H. 1994. *Naturalizing Epistemology*. Second edition. Cambridge, Mass.: MIT Press.
Kurzweil, R. 2003. “The Human Machine Merger: Are We Headed for the Matrix?” URL: http://www.kurzweilai.net/articles/art0552.html?printable=1.

Lakoff, G., and M. Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Lebedev, M., and M. Nicolelis. 2006. "Brain-Machine Interfaces: Past, Present and Future." *TRENDS in Neurosciences* 29, no. 9: 536–46.

Lovett, R. 2005. "Coffee: The Demon Drink?" URL: http://www.newscientist.com/article/mg18725181.700.

McGinn, C. 1989. "Can We Solve the Mind-Body Problem?" *Mind* 98, no. 39: 349–66. URL: http://progdata.umflint.edu/Backup/DUNLOP/PHL483/CP4%20McGinn.pdf.

———. 1991. "Consciousness and Content." In *Mind and Common Sense: Philosophical Essays on Commonsense Psychology*, edited by R. Bogdan. Cambridge: Cambridge University Press.

———. 1993. *Problems in Philosophy: The Limits of Inquiry*. New York: Wiley-Blackwell.

———. 1994. "The Problem of Philosophy." *Philosophical Studies* 76, nos. 2–3: 133–56.

Modis, Theodore. 2002. "Forecasting the Growth of Complexity and Change." *Technological Forecasting and Social Change* 69, no. 4: 377–404.

Moreno, J. 2006. *Mind Wars: Brain Research and National Defense*. New York: Dana Press.

Muller, U., N. Steffenhagen, R. Regenthal, and P. Bublak. 2004. "Effects of Modafinil of Working Memory Processes in Humans." *Psychopharmacology* 177, nos. 1–2: 161–69.

Naam, R. 2005. *More Than Human: Embracing the Promise of Biological Enhancement*. New York: Broadway.

Nagel, T. 1989. *The View from Nowhere*. Oxford: Oxford University Press.

Odling-Smee, F., K. Laland, and M. Feldman. 2003. *Niche Construction: The Neglected Process in Evolution*. Princeton: Princeton University Press.

Orca, S. 2009. "Will We Eventually Upload Our Minds?" URL: http://www.hplusmagazine.com/articles/neuro/will-we-eventually-upload-our-minds.

Poole, D., A. Mackworth, and R. Goebel. 1998. *Computational Intelligence: A Logical Approach*. Oxford: Oxford University Press. URL: http://www.cs.ubc.ca/~poole/ci.html.

Putnam, H. 1992. "Why Is a Philosopher?" In *Realism with a Human Face*. Cambridge, Mass.: Harvard University Press.

Rovane, C. 1994. "Comment on McGinn's 'The Problem of Philosophy.'" *Philosophical Studies* 76, nos. 2–3: 157–68.

Russell, Bertrand. 1936. *The Problems of Philosophy*. New York: Plain Label Books.

Sandel, M. 2004. "The Case Against Perfection." *Atlantic Monthly* 293, no. 3: 51–62.

Schneider, S. 2010. *Science Fiction and Philosophy: From Time Travel to Superintelligence*. New York: Wiley-Blackwell.

Searle, J. 1999. "The Future of Philosophy." *Philosophical Transactions of the Royal Society B*. 354, no. 1392: 2069–80.

Sellars, W. 1962. "Philosophy and the Scientific Image of Man." In *Frontiers of Science and Philosophy*, edited by Roberty Colodny, 1–40. Pittsburgh: University of Pittsburgh Press.

Strout, J. 1997. "Uploading by the Microtome Procedure." URL: http://www.ibiblio.org/jstrout/uploading/microtome.html.

———. 2002. "Nondestructive Uploading Procedures." URL: http://www.ibiblio.org/jstrout/uploading/nondestructive.html.

Tang, B. 2006. "Molecular Genetic Determinants of Human Brain Size." *Biochemical and Biophysical Research Communications* 345, no. 3: 911–16.

Tang, Y., E. Shimizu, G. Dube, C. Rampon, G. Kerchner, M. Zhuo, G. Liu, and J. Tsien. 1999. "Genetic Enhancement of Learning and Memory in Mice." *Nature* 401: 63–69.

Tsien, J. 2000. "Building a Brainier Mouse." *Scientific America*. URL: http://www.siprep.org/faculty/dlang/Eng%20475%20Homework%20 2005-06/documents/BuildingaBrainierMouse.pdf.

Verdoux, P. 2010a. "Technology and Our Epistemic Situation: What Ought Our Priorities to Be?" *Foresight* 13, no. 1: 46–57.

———. 2010b. "Human Evolution and Technology: From Prosimian to Posthuman." *Institute for Ethics and Emerging Technologies*. URL: http://ieet.org/index.php/IEET/more/verdoux20100416/.

———. 2010c. "Overcoming the 'Adaptive Lag' Through Person-Engineering." *Institute for Ethics and Emerging Technologies*. URL: http://ieet.org/index.php/IEET/more/verdoux20100819.

———. 2010d. "(Post)Human-Technology Relations." *Institute for Ethics and Emerging Technologies*. URL: http://ieet.org/index.php/IEET/more/verdoux20100329/.

———. 2010e. "The Plastic Elastic Cyborg." *Institute for Ethics and Emerging Technologies*. URL: http://ieet.org/index.php/IEET/more/verdoux20100428/.

———. 2010f. "What Are 'Biological Limitations' Anyway?" *Institute for Ethics and Emerging Technologies*. URL: http://ieet.org/index.php/IEET/more/verdoux20100218/.

———. 2010g. "Transhumanism, Progress and the Future." *Journal of Evolution and Technology* 20, no. 2: 49–69.

———. 2010h. "Risk Mysterianism and Cognitive Boosters." *Journal of Future Studies* 15, no. 1: 1–20.

Walker, M. 2002. "Prolegomena to Any Future Philosophy." *Journal of Evolution and Technology*. URL: http://jetpress.org/volume10/prolegomena.html#_ednref38.

———. 2008. "Cognitive Enhancement and the Identity Objection." *Journal of Evolution and Technology*. URL: http://jetpress.org/v18/walker.htm.

Wang, Y., D. Liu, and Y. Wang. 2003. "Discovering the Capacity of Human Memory." *Brain and Mind* 4, no. 2: 189–98.

White, N. 1998. "Cognitive Enhancement: An Everyday Event?" *International Journal of Psychology* 33, no. 2: 95–105.

Williams, C. 1997. *Terminus Brain: The Environmental Threat to Human Intelligence*. New York: Cassell.

Winner, Langdon. 1977. *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought*. Cambridge, Mass.: MIT Press.

Wittgenstein, L. 2010. *Tractatus Logico-Philosophicus*. New York: CreateSpace.

Wright, Ronald. 2004. *A Short History of Progress*. New York: De Capo Press.