AI & SOCIETY (2019) 34:321–330
https://doi.org/10.1007/s00146-018-0841-9

ORIGINAL ARTICLE

Robot as the "mechanical other": transcending karmic dilemma

Min-Sun Kim[1]

Received: 10 October 2017 / Accepted: 3 April 2018 / Published online: 17 April 2018

Abstract
As the artificial intelligence of computers grows ever-more sophisticated and continues to surpass the capacities of human minds in many ways, people are forced to question alleged ontological categories that separate humans from machines. As we are entering the world which is populated by non-enhanced and enhanced humans, cyborgs, robots, androids, avatars, and clones among them, the desire for evolutionary mastery of the natural world has taken on the two main directions: merging with machines in (1) disembodied forms or (2) embodied forms. As a path to breaking past the discontinuity between humans and machines and enter into a world beyond the "fourth" discontinuity, machines are viewed as an evolutionary step toward the "perfection" or "immortality" of humans. However, this popular, instrumental views of machines, stemming from the existential death anxiety and the hope for transcending mortality, reveals the karmic dilemma of desiring or grasping something. We shall discuss the possibility that machines can present, ultimately, a revolutionary step rather than an evolutionary step toward understanding "who we are." The path toward a continuity with machines lies not in our desire for merging with the robots, but in recognizing the arbitrary nature of all such identity categories. This radical understanding of the self-identity can be seen as a facet of enlightened experience.

Keywords Robot · Human–machine · Self-identity · Evolution · Revolution

1 Introduction

> "The time has come," the walrus said, "to talk of many things: Of shoes and ships—and sealing wax—of cabbages and kings".

—Lewis Carroll, Alice's Adventures in Wonderland & Through the Looking-Glass.

Robots are clearly a present reality, not a fantasy. As humans come increasingly to view themselves as living in a mechanical or simulated universe, automata now take the shape of artificial intelligence machines. Modern machines are evolving from a more-or-less neutral channel of message transfer and instrument of human interaction to the position of participant in communicative exchanges (Gunkel 2012a; Kim et al. 2009, 2011). Margaret Somerville (2006) argues that the species *Homo sapiens* is evolving into *Techno sapiens* as we project our abilities out into our technology at an accelerating rate. It is time, now, to reflect on the revolutionary, rather than evolutionary, journey from the beginnings of evolution to the latest computers/robots. According to Mazlish (1993), biogenetics appears to place us between the hope, for some people, of producing a superhuman, and the fear, by many, of conceiving monsters.

As one of the most enchanting of ancient and modern machines, robots are the locus of a variety of interests such as economic, military, household, scientific, and spiritual, and so on. So far, most humans generally seem to deny that such artificially intelligent and autonomous entities will actually emerge, declaring that if they do, they will be vastly inferior to humans, and thus not worthy of human respect or spiritual concern (Dator 1989). Likewise, some have taken up the Frankenstein-like fears of the "monster" machines, now evolving, and threatening human survival (Butler 1923). According to Stephen Hawking, "As the Hollywood blockbuster *Transcendence* debuts this weekend with Johnny Depp, Morgan Freeman and clashing visions for the future of humanity, it is tempting to dismiss the notion of highly intelligent machines as mere science fiction. But this would be a mistake, and potentially our worst mistake ever."

✉ Min-Sun Kim
kmin@hawaii.edu

[1] Department of Communicology, University of Hawaii at Manoa, Honolulu, HI 96822, USA

Variously termed machines, automata, or androids, AI, or robots, they all pose the same ontological question: how do they differ from humans, or, what does it mean to be human? How will human behavior evolve with the inclusion of intelligent robotics? Rodney Brooks (2002) has stated that he will know that he is getting close to his goal of building an intelligent robot when his graduate students feel guilty about turning it off. Titles of some books relating to machines in general can be a foretaste of the pleasures, or horrors, to come: *Are Computers Alive?—Evolution and New Life Forms; Superintelligence: Paths, Dangers, Strategies; Artificial Life; Internet Afterlife: Virtual Salvation in the twenty-first Century; The Second Self; Our final Invention; Love and Sex with Robots: The Evolution of Human–Robot Relationships; Are You a Machine?: The Brain, the Mind, And What It Means to Be Human; Singularity Hypotheses: A Scientific and Philosophical Analysis*.

We have come a long way from the beginnings of evolution to the latest computers/machines. The world has changed in unimaginable ways. At the same time, videogames are becoming more and more sophisticated and in the future there may be simulations of conscious entities living inside machines. It is time to reflect on implications of the machines and (r)evolutionary journey, asking what it means to be a human and how we relate to the "Other," such as animals and, more recently, to machines. As we are entering the world which is populated by non-enhanced and enhanced humans, cyborgs, robots, androids, avatars, and clones among them, the transcendence of the discontinuity with machines has taken on the theme of instrumental views of machines as an advanced evolutionary species. In this paper, I will discuss the possibility that machines may present, ultimately, a revolutionary step rather than an evolutionary step toward understanding "who we are."

2 Coming of a robotic age: what does it mean to be a "human"?

As Marshall McLuhan (1964) argued, "We first shape the tools and thereafter our tools shape us." What does it mean to be a human in an era wherein human conjoins with machines, biology with technology? We are in the midst of a technological revolution. Dator and Seo (2004) suggested that the world may be moving beyond both an "industrial" and an "information" society (based so heavily upon the printed word) into what we call "a dream society of icons and aesthetic experience." In such a world, governance, war and religion disappear as major elements as individuals spend more and more time in virtual interaction with people (and, increasingly, also with artificial intelligences) of their own choosing without regard to physical, geographic proximity or to the conventions of "real time" compared to free-flowing "dream time." While this becomes preferred future of some people, the interrelationship of humans with increasingly intelligent, autonomous, or simulated artificial entities, variously called robots, artificial intelligence, cyborgs, avatars, or artilects pose significant emotional and spiritual implications for human race. With the reality of intelligent machines, we are on the edge of the new conception of humanity.

According to Pribram (2013), the mechanical revolution has many of the earmarks of the Copernican revolution of earlier times. "That revolution inaugurated a series of scientific breakthroughs such as those by Galileo, Newton, Darwin and Freud. Though often unintended, all of these and related contributions shifted humans from center stage to over more peripheral players in the scientific scenario" (p. 8). For some people rapid technological progress and the prospect of intelligent machines challenge the sanctity of humankind, and may cause anxiety due to the potential to cause damage to "uniquely human" capacities—physical, mental, moral, and spiritual.

The riddle of human existence and human consciousness has been pondered throughout the history. Today, a radical understanding of human nature depends on the development of intelligent machines and "artificial" environments (i.e., simulations) created by them. Although modern science—like religion, philosophy, and literature throughout Western history—has itself perpetuated myths about human and non-human animals alike, in recent decades, there have been dramatic breakthroughs in science that have advanced understanding of the human nature (Mazlish 1993). Sophisticated computer technologies and artificial life are revolutionizing our self-identity through more accurate glimpses into the human relationship with machines.

3 Discontinuity between humans and non-humans: coping with animality and mortality

A widely accepted belief is that robots are nonhumans and should be positioned outside all humans as a human-made technological artifact (see Kim and Kim 2013). When technologies are reduced to mere tools, there is little space left for reflecting on the ways in which the creation of 'compuhumans' will affect our experience of ourselves and the world. Some humans insist on their separateness and superiority in regard to machines (as well as other animals), viewing them as a threatening new "species" (Mazlish 1993). Sometimes, humans describe the evolutionary form of new species as "the fearful Other," or "Subhuman Other," showing racist fears toward robots and A.I. (see Heo and Kim 2013; Kim and Kim 2013).

Graphically rendered robot characters are coexisting with synthetic human characters of different races in the myriads of computer interfaces. As the artificial intelligence of computers grows ever more sophisticated and continues to surpass the capacities of human minds in many ways, people are forced to question alleged ontological divide, the one separating humans from machines. To what extent are humans unique or distinct from everything else in the world? Psychohistorian Bruce Mazlish (1993) discusses that in the historic "smashings" of the ego, humans are placed on a continuous spectrum in relation to the universe, to the rest of the animal kingdom, and to themselves. In his book, *The Fourth Discontinuity* (1993), Bruce Mazlish identifies four ruptures in the Western medieval picture of reality brought about by dynamic changes in the modern world.

The *first* discontinuity started with the Copernican revolution in the sixteenth century which destabilized the geocentric worldview. According to Mazlish (1993), Sigmund Freud suggested that achievements by Copernicus helped remove the first traces of anthropomorphism from science. Freud (1917, p. 1916) wrote that "our earth was not the center of the universe but only a tiny fragment of a cosmic system of scarcely imaginable vastness." In place of the worldview that situated the Earth at the epicenter of the universe, Copernicus, and subsequently Galileo in the seventeenth century, argued that the sun occupies the center of the universe and the Earth revolves around the sun. Under the spell of the Ptolemy and medieval cosmology, human beings had to confront the fact that their planet is not the physical center of the universe (Mazlish 1993). Not only did this fact contradict official Church dogma, the spatial decentering entailed a psychological decentering, moving the Earth and possibly humanity itself from the center of the picture to the margins (Best 2009).

The second smashing of the naïve self-love of humans was opened up in 1859, when Darwin published *The Origin of Species*. According to Freud (1917), Darwin revealed that the exceptional discontinuity we perceived between ourselves and other animals or plants was equally illusionary. In *The Origin of Species*, Darwin wrote that "We are one continuous life, one evolution." "Man in his arrogance thinks himself a great work, worthy the interposition of a deity. [Yet it is] more humble and, I believe, true to consider him created from animals." Science subsequently has confirmed that animals lived for billions of years before humans, and all life evolves in a continuum from the same primordial conditions (Mazlish 1993).

Western culture had to confront the facts and consequences of a *third* discontinuity opened by the theory of the unconscious mind by Sigmund Freud. Freud (1917) claimed that psychoanalysis seeks to prove to the ego that it is not even master in its own house, but must contend itself with scanty information of what is going on unconsciously in the mind. Against the Christian/Cartesian view of the self as governed by a rational command center and the body as a temporary housing for the immortal soul, Freud (1917) demonstrated that rationality and conscious thought are products of the unconscious realm of existence governed by primordial instincts, desires, drives, and the sexual and violent urges of the Id. Likewise, in 1917, Freud wrote "A Difficulty in the Path of Psychoanalysis," concerning the three great shocks to man's [sic] ego, as described above.

Quoting Freud's discussion about the cosmological, biological, and psychological blows to human pride, Mazlish (1993) goes on to reveal yet another discontinuity that we have yet to recognize as illusory, *the boundary between humans and machines*. Mazlish notes, a *fourth* discontinuity surface during the mid-twentieth century, with the rapid development of computer technologies and artificial intelligence. Those shocks administered by Copernicus, Darwin, and Freud compelled humans to confront their separation from the cosmos, their animal origins, and the primacy of their subconscious being. Again, humans are forced to reconsider their relation to thinking machines.

For much of the past 500 years, Western scientists and philosophers have taken it for granted that human beings are exceptional creatures, not simply distinct from other animals, but superior to them, because of our possession of reason and consciousness, language and morality. But according to Malik (2001), rather than sustaining this tradition of 'exceptionalism', some recent developments in evolutionary biology, advances in genetics, neuroscience and artificial intelligence pose new challenges to long-standing ideas of human distinctiveness. Humans are confronting the reality that their own evolution is inextricably interwoven with the use and development of tools.

As the artificial intelligence of computers grows evermore sophisticated and continues to surpass the capacities of human minds in many ways, people are forced to question yet another alleged ontological divide, the one separating humans from machines. According to Best (2009), machines are no longer mechanisms in traditional sense, since they are ever more closely approximating the biological operations of the human brain through neural nets, parallel processing, evolutionary hardware, and the like. Moreover, when human beings begin to merge intimately with their machines, fusing flesh with steel and silicon chips, human identity comes into question.

Artificial autonomous agents are the next stage in the development of technology. In that sense, the normal concerns about the impact of technological development apply, but in another sense they represent a "profound discontinuity" (Poole and MacWorth 2010, p. 16). As autonomous agents perceive, decide, and act on their own, there is a radical and qualitative change in our image of technology and

humans are forced to reevaluate their place in this mechanical world.

According to Gunkel (2012b), when nonhuman otherness for animals is increasingly recognized as a legitimate moral subject, its other, the machine, remains conspicuously absent and marginalized. Like many before him, Malik (2001) posits the human as different from all other creatures and so implies a notional boundary between the human and nonhuman domains. The tradition of exceptionalism had led humans to presume they were physically and mentally distinct, bounded by and from the world around them. However, it seems we are entering a new phase of socio-cultural development, one that is sometimes called 'posthuman', in which humans are increasingly seen as less distinct from animals, machines and the environment. These developments make it more difficult to maintain the notional boundaries that for so long have held humans apart from the world and from each other (Pepperell 2005).

4 Machines as evolutionary heir: instrumental views of robo-sapiens

The terms such as 'posthumanism,' along with 'transhumanism' has been used in various different ways. Francis Fukuyama (2002) in his book, *Our Posthuman Future*, pronounced "the end of history" and argued that as a result of biomedical advances, we are facing the possibility of a future in which our humanity itself will be altered beyond recognition. Thus, the posthuman is the biotechnologically mutated "non-human"—a creature that remains biologically encased but divorced from its natural biological origin. For others, the posthuman is the technologically encased successor to the soon-to-be-obsolete biological human—a cyborg-entity inhabiting data-space, enjoying a computationally generated consciousness unconstrained by the physics of corporeal existence and the doom of mortality.

As humans create intelligent computers and robots, humans become ever-more like cyborgs by incorporating technology into the human body. One of the most frequently discussed topics in posthumanist and transhumanist circles is life extension, which normally denotes the indefinite prolongation of life through chemical, cryogenic or other technological means (Pepperell 2005). The expectations bound up in technologies like xenotransplantation, artificial consciousness and intelligence, synthetic replication, biotechnical integration and cloning signifies the hope of liberation from death. As we may expect to share organs with pigs, sport prosthetic limbs, find ourselves digitally extended into the world by a telepresence system, or confront identical copies of ourselves, the potential transcendence from the human predicament seems more and more graspable.

Humans are on the threshold of decisively breaking past the discontinuity between themselves and machines. Going beyond the mere smoothing with machines in general, humans are yearning take on some of the superior attributes of "intelligent machines" or to rise above their own animal nature. That is, human nature is considered to evolve in intimate connection with humanity's creation of intelligent machines. The typical modes of transcendence of the fourth discontinuity have taken on evolutionary perspectives with the computer-robot as a possible "evolutionary" form. Darwin himself says that in the last paragraph of *The Descent of Man*, the theory of evolution offers Man "hope for a still higher destiny in the distant future" (p. 920). The higher destiny can mean "being angelic," or "being super human." Samuel Butler (1923) raises the idea of a new species—"machia-species." Similarly, according to Mazlish, something like a new species will eventually emerge—*homo comboticus*—that will compete with and very likely replace (or convert) most of the human types that have existed before about 1970; that is, precomputer Man. One can distinguish phases in that evolution: prosthetic extensions; prosthetics joining of human and machine, culminating in human-cum-computer; and coming into being, at the hands of humans, of a new type of species, the thinking machine.

Throughout the history, humans have struggled intensely with the question of their being: animal or angelic, human or machine? The Utopian quest for technological salvation is inseparable from evolving human nature. The typical modes of transcendence with evolutionary views of machines have taken on two main directions: (1) toward disembodied machines, and (2) toward embodied machines. Put it differently, there is the posthumanism of disembodiment, which wants liberation from the encumbering limitations of the physical realm. Then there is the posthumanism of embodiment, which recognizes the essential nature of the bodily dimension with extended corporeal realm of sensory perceptions in merging with machines.

4.1 Disembodied evolution

While many resist the implosion of biology and technology, a bold cadre of technophiles, visionaries, futurists, AI researchers, and transhumanists embrace it as the next and inevitable stage in human evolution. In one version, humans will soon be able to create "spiritual machines" (Kurzweil 1999) or "mind children" (Moravec 1999) that constitute a new posthuman species far superior to our current carbon-based model.

Back in 1972, Hubert Dreyfus's (1972) manifesto ("What Computers Can't Do") on the inherent inability of disembodied machines to mimic higher mental functions caused an uproar in the artificial intelligence community. The typical views on disembodied evolution of humans via technological

means, however, do not focus on AI research and engineering principles. For instance, Michael Foucault (2006) speak of a utopian ideal of the relation of the body to technology. He defines Utopia in his short essay, Utopian Body as, ...a place outside all places but it is a place where I will have a body without a body that will be beautiful, limpid, transparent, luminous, speedy, colossal in its power, infinite in duration. Untethered, invisible, protected-always transfigured. It may well be that the first utopia, the one most deeply rooted in the hearts of men, is precisely the utopia of an incorporeal body (p. 229).

Moravec (1999) envisions humanity moving to a higher state of being and attaining immortality by merging their minds with computers. Far more than theories of evolution, these techno-utopias represent neo-Cartesian assumptions that mind is substance and body is an accidental trait as well as secular manifestations of the Christian quest for immortality, promoting the process of merging with our machines, of creating "man-machines".

What has been emphasized is that humans are evolutionary beings. Angels were a marker on the Christian way to human perfection. Machines took on the same quality for more secularly minded humans. Embodied in the idea of progress, one way is to lead humans into a mechanical paradise in which they were perfectable, thus leaving the human as a purely spiritual creature (Mazlish 1993).

Robert Jastrow (1983), in his book *The Enchanted Loom: Mind in the Universe*, talks of the human brain, ensconced in a computer, becomes liberated from the weaknesses of mortal flesh, creating a new form of existenc. Housed in indestructible lattices of silicon, and no longer constrained in the span of its years by the life and death cycle of a biological organism, such an existence could live forever. According to this view, we are on a continuum with the machines we have created, transcending the weaknesses of mortal flesh. In the disembodied evolution, it is humans who will evolve into advanced "robots." Their belief is that with futuristic technologies being developed in multiple fields, human intelligence may eventually be able to biological tissue and be able to move freely across boundaries that cannot support flesh and blood.

Hans Moravec (1999) predicts that machines will attain human levels of intelligence by the year 2040, and that by 2050, they will surpass us. Far from railing against a future in which machines rule the world, Moravec takes the view that intelligent robots will actually be our evolutionary heirs. According to him, intelligent machines will learn human skills, and share human goals and values, can be viewed as children of human minds. And since they are our children, we will want them to outdistance us. In fact, in a bid for immortality, many of our descendants will choose to transform into "ex humans," as they upload themselves into advanced computers. The robotics researcher Hans Moravec has been advocating the notion that human consciousness is essentially a process of abstract symbolic manipulation; one that in the future could be simulated in a disembodied computational medium as bodiless mind.

Similarly, Ray Kurzweil predicted, in his 1999 book *The Age of Spiritual Machines*, that non-biological intelligence will vastly exceed the collective brainpower of *Homo sapiens* within this century, and that the human race will voluntarily merge with technology, so that by 2100 there basically will not be any traditional humans left. The disavowal of the unique and distinct human tends to lead to the abstracted, dislocated existence touted by some posthumanists. According to the disembodied version of evolution, it will be a being that, if it ceases to be human at all, will not be abandoned as a redundant shell. The brain will finally be free to travel among the stars. Through the mechanically embodied existence we may find salvation from the limits of bounded experience and human finitude (see Kurzweil 2005). Lombardo (2017) similarly proposes the notion of the "World Brain" and "World Mind" as the overarching global expression of the evolving human–technology integration. He argues that this psychophysical system will enhance and enrich the capacities of both individual and collective cognition and serve as a potential starting point toward the evolution of a cosmic brain and cosmic mind.

Although some futurists are given to the optimism.

4.2 Embodied evolution

While taking a posthumanist stance, Pepperell (2005) criticizes the widespread tendency amongst advocates of posthumanism toward 'disembodiment'—the proposed distillation of human essence into some immaterial form. He does not believe that humans or human experience can be reduced to an essence—digital or otherwise—free from the contingencies of corporeality. According to Pepperell (2005), "Although a highly attractive scenario for those wishing to mechanically emulate human beings, this kind of dualism overlooks some crucial aspects of human existence."

Since human beings are conditioned by physical and biological constraints, without those corporeal dimensions, human experience would have a very different meaning, or perhaps no meaning at all (Pepperell 2005). For instance, our sense of subjectivity—the knowledge we have of ourselves as sentient beings—does not seem to be wired into our brains in the form of some pre-given 'program'. Rather it develops as humans interact dynamically with the world and other sentient beings who also occupy mobile bodies, giving rise to what the philosopher Edmund Husserl called 'intersubjectivity.' Likewise, our worldly experiences are physically incarnate as much as they are mentally experienced; they arise from, and gain their meaning from, the extended corporeal and flesh realm of sensory being (Pepperell 2005).

Support for an embodied interpretation of posthumanism, and recognition of its historical importance, can also be found in recent critical commentaries in the field. In "How We Became Posthuman," Hayles (1999) criticizes the tendency found in transhumanism in regards to human future as becoming technologically disembodied. She argues for the importance of putting embodiment back in the picture. According to her, embodiment makes clear that thought is a much broader cognitive function depending for its specificities on the embodied form enacting it.

According to Pepperell (2003), we need to come to a different understanding of our human predicament from some of the posthuman advocates whose concern with extending life by migrating consciousness from brains to machines. Those who advocate the embodied version of merger with machines argue for extended *experience*, and an experience based on the contingencies of embodied existence at that. The real posthuman, then, is properly conceived not as an abstract flow of bits or information but as a radically extended and *embodied* being whose experience, which is potentially boundless.

The two main understandings of technology (either disembodied and embodied) in an evolutionary context is based on the instrumental views toward robots—a means to an end. Heidegger (1954) warns of the inherent dangers in such advantages, however, because the more sophisticated our technology becomes, the more the instrumental conception of technology challenges the right relation to technology.

This popular, instrumental views of machines, stemming from the existential death anxiety and the hope for immortality, reveals the karmic dilemma of desiring or grasping something. To cut through the allure of machines promising perfection and "salvation," we will ponder a radically different notion of personal identity.

5 Robots as revolutionary mirrors for humans

> "Immortality is a condition fairly difficult to come by, though the longing for it is in one way or another almost universal and considered by some, therefore, to be beneath one's dignity."
>
> —James Lord.

The theme of death is a generalized sense of the threat posed by the perceived artificial and lifeless nature of the mechanical world. Ironically, some regard the new mechanization of humans and society and see not death, but life ("triumph over death"). Mechanization means denying organic decay, and in its most extreme version, giving mankind a secular form of eternal life. Symbolic of the mixed emotions of fear and hope surrounding the emerging mechanical world, humans are confronted with, for better or worse, the possible scenario of a quantum jump in the evolution of the human species into a new "species," a mechanical "being" (Mazlish 1993).

Historically, the emergence of new technologies provides the base for profound changes in the structure of self. There is understandable anxiety aroused by the prospect of the human as either technologically de-materialized or technologically displaced—that is, carrying on in some new form or not carrying on at all. However, expectations about the victories over the body have not engendered the forms of realization, or fostered new sensibilities regarding the understanding of self.

In Western cultures, the bridging of the boundaries between humans and machines is becoming increasingly visible primarily through the potential technological victory over death. In other cultures, the distinction or gap between humans and everything else (including machines) was never very strong, for instance, for the Achuar Indians nature and society are part of continuum rather than independent spheres (Descola 1994). Bostrom (2004) argues that current evidence does not warrant any great confidence in the belief that the default course of future human evolution via technological means points in a desirable direction. In particular, he proposes dystopian scenarios in which evolutionary competition leads to the extinction of the life forms humans regard as valuable.

According to Mazlish (1993), "it is exactly the most characteristic traits of the human condition—for example, fear of death, loathing of the body, …the desire for evolutionary mastery of the natural world, comprise the fundamental forces that increasingly impel humans toward the "merging" with machines" (p. 219). For some, the machine promises eternal life. In regard to the limit of death, biocentrism—a new theory of everything—tells us death may not be the terminal event we think. Immortality does not mean a perpetual existence in time, but resides outside of time altogether (Lanza and Berman 2010).

5.1 You and I, mere holograms

> "I wonder if I've been changed in the night. Let me think. Was I the same when I got up this morning? I almost think I can remember feeling a little different. But if I'm not the same, the next question is 'Who in the world am I?' Ah, that's the great puzzle!"
>
> —Lewis Carroll, Alice in Wonderland.

Typically, the issue of human mortality is intimately connected with the view that machines represent an evolutionary step toward the "perfection" or "immortality" of humans. This popular position stemming from death anxiety is derived from the self-other cognitive split. The

problematic grasping or quest for "immortality" via technology can viewed as holding on to a separate, self-centered existence. Through growth to a higher level, through rebirth to a new attitude and a new way of being, one can conceive that there may be no such thing as personal identity. In a way, the no-self theory is not a theory about the self at all. It is rather a rejection of all such theories (on the existence of self) as inherently untenable (Giles 1993). Similarly, in Zen experience, word loses all meaning since it's all there is; Zen declares that words are words and no more. No matter what verbal space you try to enclose Zen in, it resists, and spills over (Hofstadter 1980).

All societies, albeit in vastly different ways, are systems of illusion and attachment, and social customs of all kinds tend to perpetuate ego-based assumptions about the identity and continuity of the self. Perhaps, there is nothing behind the curtain and there is no spoon. However, counterintuitive this proposition ("The Theory of No-Self") may be, identity (or personhood) is the deepest and most universal social illusion; it is also humanity's number one problem, the thing we most need to figure out to get on with self-(r)evolution. One of the important implications of this approach, not only for AI but for a general understanding of the world, is that it bridges the Human versus Others cognitive split.

The theory of no-self (Giles 1993) resonates with some ancient philosophical traditions. Buddhist thinkers have long spoken of the 'non-self' or anatta whereby the apparent distinctions giving rise to a sense of unique, distinct self are abrogated. Similarly, Humphreys (1985) wrote that the third Sign of Being is anatta, which literally means that no 'compounded thing' has an atta (Sanskrit: atman). The Buddha taught that in none of the constituents of the personality, the physical body, feelings, reactions, various mental attributes and discriminative consciousness is there a permanent element which distinguishes that man from any other.

Among all Buddhist concepts, the notion of no-self (Sanskrit: anatman) seems to be the single most important concept within the Buddhist tradition and, at the same time, the least intelligible one to most (Kopf 2001). Recently, however, the Buddhist theory of no-self finds its conceptual equivalent in the outright rejection of an underlying personal identity or self. As one of the most important no-self theorist, David Hume (1896) was the first Western philosopher to unmask the confusions attending our idea of personal identity and subsequently to reject the idea as a fiction. In "A Treatise of Human Nature" (1888), Hume points out that although some philosophers believe we are continuously aware of something we call the self, when we look to our experience there is nothing to substantiate this belief. What we experience, rather, is a continuous flow of perceptions that replace one another in rapid succession. In other words, there is no self to be found (Giles 1993).

Similarly, Dogen Kigen's Zen Buddhism developed a phenomenology of no-self which suggests a conceptual strategy to respond to the questions of personal identity and theorizing a selfless self after the refutation of the notion of personal identity and after the loss of an enduring self. Ultimately, Dogen suggests a radical rethinking and reversal of the individualistic conception of self. For the self is not an enduring individual which "has" experiences and psychic states, but constitutes an experience or, more appropriately, an awareness event (Kopf 2001).

Paul Virillo (1994), in "The Vision Machine," wrote that under post-modernity the individual's relation to reality is collapsing. The Buddhist doctrine of no-self (Skrt.: anatman) proposes a way to theorize selfhood in this increasingly mechanical world. No doubt mirroring the claims of the new technologies for prolonging biological life that are hyped daily in the latest long-awaited miracle from science, the techniques to extend life stimulate rather than alleviate the natural repressed anxiety over death. Kasulis (1981) remarks that "to be Dogen's primordial person is to be essentially no person, while simultaneously being the personal act appropriate to the occasion" (p. 154). Human experiences is not appropriated by a pre-existing person-over-time, abut, on the contrary the enduring person-over-time is constructed as abstract personal identity. Dogen, however, believes that this construction does not comprise a necessary condition for responsibility; on the contrary, responsibility is engendered by an unself-conscious attending to the *demand of the present*.

Kopf (2001, p. 259) uses the term "decentralization" to indicate a gradually decreasing importance of self as the center of human activity. With the challenge to, and the undermining of, the essentialist paradigm, which began with Hume, "The proposition that the self is an illusion, a socially constructed reality—that there are quite different ways of thinking about personal identity—seems to contradict plain common sense. And even people who accept the idea in the abstract do not necessarily get it in a way that makes a difference to how they experience daily life. That other, much older enlightenment project—the one that we associate with the Buddhists and the Sufis—was also built around a radically different notion of personal identity, a quest for liberation from the ego." (Kopf 2001, p. 218).

According to Anderson (2003), speculating about "no-self" from the unquestioned standpoint of self may make it seem implausible, even bizarre, ridiculous, or downright self-contradictory. He acknowledges that describing identity as a problem is radically different from our usually tendency, which is to describe it as a fundamental human need, only problematic when absent. People generally assume that to be "real" people we need to have a full list of labels signifying nationality, racial heritage, religious affiliation, political beliefs, occupation, position in society, gender and sexual

orientation, among others. "With all these in order we are in good shape. Without them we are nobody. And if we are somewhere in between, uncertain about which labels we should be wearing and when, we are diagnosed as sufferers from alienation, anomie, identity crisis, or other social diseases identified by modern theorists" (Anderson 2003, p. 170).

There is an old discourse that revolves around the proposition that self is an illusion and has no existence except as an abstract concept (Anderson 2003). Varela et al. (1991), in their book, *The Embodied Mind*, suggest that Western science should begin to reformulate its cognitive notions of the ego-self. The authors overview numerous Western cognitive theories—most of which markedly exclude an examination of phenomenological experience- and conclude that we should bear in mind the discoveries of Buddhist tradition if we are to understand what we call "self" is, according to Buddhist practices, no more than an epiphenomenon, a fictional construct that results from a continuous pattern of "grasping." "Constantly one thinks, feels, and acts," the authors write, as though one had a self to protect and preserve. The slightest encroachment on the self's territory (a splinter in the finger, a noisy neighbor) arouses fear and danger. The slightest hope of self-enhancement (gain, praise, fame, pleasure) arouses greed and grasping.... Such impulses are instinctual, automatic, pervasive, and powerful (Varela et al. 1991).

Desire and grasping forms the "karmic process" stemming from the "karmic habits," which can be defined as the autonomous and unthinking repetition of an action, repeating a conditioned action (Kopf 2001). Likewise, Lusthaus (1989) remarks:

> Karmic actions, whether cognitive in the broad sense or strictly mental, therefore are always intentional, and the karmic dilemma is a dilemma of intentionality. 'Intentional' implies 'desire', intending or desiring something" (p. 168).

Yet when we recognize this grasping for what it is—in that it alone creates what we think of as self—we recognize too that our individual identity has in truth no solid ground, that "we instead emerge, moment to moment, from our physical interactions with the world" (Galvan 1997, p. 424).

Various authors argued that we are entering a period in which we are coming to recognize the continuities not just between humans and machines, but between all things that might previously have been held as bounded and separate. In doing so we are becoming increasingly posthuman (Pepperell 2005). In human conception of technology, the danger is the way intelligent machines are perceived. If machines are perceived in the "instrumental" mode toward human "perfection" or "immortality," then this grasping becomes karmic dilemma. Buddhism conceives of the context created by karmic activities as habit-formation, while casting the concept of *nirvana* as the cessation of karmic activities. The path towards continuity with machines lies not in our desire for merging with the robots, but in recognizing the trivial and arbitrary nature of all such identity categories. This radical understanding of the self-identity can be seen as a facet of enlightened experience.

Everyday awareness, as the first realm of experience, is characterized by a binary structure and naïve world view, which takes *appearances* at face value (Kopf 2001). The universe, as the sages in every religion teach us, may be one vast continuum. To utilize knowledge of this universal continuum, we need to think of our technology as really a way of expanding our collective mind, and engendering existential reorientation. Whereas, in the mythical land of "Oz," reality stems from the wizard's conjuring trick, in the quantum realm, Bohr (1935) argued, there is no wizard. There is "nothing" behind the curtain; all we see is the formless archetypal play of phenomena itself, a display which is empty of inherent existence and inextricably linked to our consciousness and its various operations.

The similarity to the Buddhist idea of "dependent co-arising" (also called interdependent co-origination), which is considered to be the very condition of and process by which empirical reality is constituted. Dependent co-arising is considered to be a milestone in human thought; its ever-deepening realization is one of the greatest and furthest reaching cognitive revolutions of our time. To recognize the lack of permanent entities, such as human selves, is to recognize dependent origination.

Baudrillard (1994), in his book, "Simulacra and Simulations," claimed that "the simulacrum is never that which conceals truth—it is the truth which conceals that there is none." What we experience is actually the result of data gathered about the physical world; so it cannot be that world itself. Rather, the data about the phenomenal world is collected, processed, and then integrated into a *simulation* of the world. This simulation is our experience of the world. Thus, according to Baudrillard (1994), our experience is nothing more or less than a simulation generated from data from our senses.

Most of us live in a piece of territory called "me"—a personality who seems to live inside the boundaries of its skin, separate from all that surrounds it. Our minds tend to be consumed by myths and illusions of personhood. This separate personality thus involves itself in minimizing any potential discomfort or pain, and maximizing its pleasure, hoping for a happier future filled with good things and experiences. Enlightenment is noble cognition, the seeing past collective error and illusion to a hidden reality.

In Philip K. Dick's (1968) science fiction novel, *Do Androids Dream of Electric Sheep?*, Rick himself realizes, bounty hunting of replicants no longer fortifies an inherited

notion of himself as subject; indeed, it rather challenges all that he understands himself to be. Now that he has carried out the task he has been appointed to, he finds himself "defeated in some obscure way" (Galvan 1997). Yet that defeat-of an old understanding of self against world—also marks, paradoxically, Rick's triumph: his new awareness that he lives in fluid conjunction with the technologies that populate his environment. There is no human self, Rick has discovered, that is not also other, and no android other that does not partake of self. In one scene, having found in the desert what he believes to be a natural toad, Rick hurries home tell his wife, only to have her reveal to him that the toad is mechanical. But in reacting to the news, he remarks, "The electric things have their lives, too. Paltry as those lives are" (Galvan 1997).

In describing Rick's interaction with the mechanical landscape, Galvan (1997) points out Rick's realization that technology is indeed a vital part of the planetary environment. To have overlooked this reality has meant denying the basic entre-deux between self and world and denying, specifically, the established presence of diverse machines, which are materially intertwined into the lives of the novel's characters (Galvan 1997). The novel reminds us of the potential dangers of instrumental view of machines and technology. Rick, the bounty hunter of androids, at last conceives that technology always already impinges on the human subject, always already cooriginates with him. It is up to each individual to relinquish a self that has outgrown traditional human categories into the posthuman collective (Galvan 1997).

Perhaps this notion of immortality gets us closer to the strange logic of artificial life, particularly as it is being understood by the ways in which the simulated, immersive worlds of hyper-realistic graphics where robots are "technological human simulacra" (Olivier 2008, p. 30). According to Tofts (2003), "The concept of vitality, in the context of simulation, is highly ambivalent. It is not a dualist model of one thing or another, but rather it is one thing and the other, living and dead at one and the same time" (p. 59). Therefore, in the artifactual world, the beginning of life is not so much a genetic emergence, but rather the iteration or re-activation of a previous suspension. "Not really alive, artificial life plays at being alive. Never really dead, it reanimates indefinitely" (Tofts 2003, p. 59).

6 Conclusion

For Feenberg (1999), technology is the most important issue of our era. It is a major constituent of contemporary society and is intimately connected with politics, economics, culture, and all forms of social and personal life. Since the opening of modernity five centuries ago, human beings have had to confront major discontinuities which problematized their uniqueness and special status in the universe (Mazlish 1993). In quick succession, *Homo sapiens* had to overcome scientific and philosophical false dichotomies and illusions of separation from the *infinite cosmos*, the *animal world*, the *unconscious*, and *machines*.

We are immersed in an artificial world which turns fantasies into realities on a minute-by-minute basis. Automata now take the shape of artificial intelligence machines. The debate over human and machine conjures up before us the specter of a "new species." The Darwinian vision of machines as an evolutionary development after "Man" has taken various forms. The prospect of extending Darwin's theory of evolution to machines can be seen in the service of human survival. The popular position has been that we are on an evolutionary continuum, with machines as a new, and possibly advanced, species. The popular notions of either embodied or disembodied merging with machines are taking the form of the "techno-spiritual materialism."

So far, the various attempts to overcome the discontinuity with machines is based on ego-based assumptions about the identity and continuity of self. Fundamental longings for the extension of life through merging with machines are not new, but are represented in the human aspiration typical of classical mythology, the quest for immortality, the elixir of life taken up by the alchemists, and hopes for enlightenment. Longings for immortality persist even while so many of the stories we tell about them end badly. These longings constitute "karmic dilemma" (Lusthaus 1989, p. 168), or dilemma of "desiring something." Consciousness of the dreamer in this technological utopia will still be divided against itself, creating the state of techno-enlightenment, which blinds rather than enlightens. Metaphorically, one can be reminded that Harry Potter is the only known Master of Death, having gathered and mastered all three Hallows and most importantly, accepted death as old friend.

As the science and technology of AI continues to develop, there will be profound ethical, psychological, social, economic, spiritual, and legal consequences for human society, our planet, and beyond. According to the projections outlined in Damien Broderick's *The Spike: Accelerating into the Unimaginable Future*, the acceleration of change is increasing so sharply that the future may not be just unknowable, but unrecognizable by "humans." Further analysis is needed to explore the implications for transformations of human consciousness. The longings for a sense of "connection" among all "alienated" beings may be illusory, since nothing is "real" (Sims 2009). Theodore Roszak (1986) notes that the "salvational longings ... entwine themselves around new technology" (p. 45). The revolutionary views of "no-self" may cut through the instrumental attitudes toward intelligent machines as the path to immortality and open the path to being at ease in the present.

References

Anderson W (2003) The next enlightenment: integrating east and west in a new vision of human. St. Martin's Press, New York

Baudrillard J (1994) Simulacra and Simulation. Tr. Sheila Faria Glaser. The University of Michigan Press, Ann Arbor

Best S (2009) Minding the animals: ethology and the obsolescence of left humanism. Int J Incl Democr 5(2):1–26

Bohr (1935) Can quantum mechanical description of physical reality be considered complete? Phys Rev 48:696–702

Bostrom N (2004) The future of human evolution. In: Tandy C (ed) Death and anti-death: 200 years after Kant, 50 years after turing. Ria University Press, Palo Alto

Brooks RA (2002) Flesh and machines: how robots will change us. Pantheon Books, New York

Butler S (1923) Darwin among the machines. In: In a first year in Canterbury settlement and other early essays, vols 20. Shrewsbury Edition of the Works of Samuel Butler, London

Dator J (1989) "What do 'You' do when your robot bows as your clone enters holographic MTV?" for the conference, "humans-cultures-futures," Tenri City, Japan, August 30–31, 1988, published in *Futures*

Dator J, Seo Y (2004) Korea as the wave of a future: the emerging dream society of icons and aesthetic experience. J Futures Stud 9:31–44

Descola P (1994) In the society of nature: a native ecology in Amazonia (Nora Scott, Trans.). Cambridge University Press, Cambridge

Dick PK (1968) Do androids dream of electric sheep? Del Rey, New York

Dreyfus H (1972) What computers can't do. MIT Press, New York

Feenberg A (1999) Questioning technology. Routledge, London

Foucault M (2006) "Utopian body". Sensorium: embodied experience, technology and contemporary art. MIT, Cambridge

Freud S (1917) A difficulty in the path of psychoanalysis. SE 17:137–144

Fukuyama F (2002) Our Posthuman future: consequences of the biotechnology revolution. Farrar, Straus and Giroux, New York

Galvan J (1997) Entering the posthuman collective in Philip K. Dick's "do androids dream of electric sheep?" Sci Fict Stud 24(3):413–429

Giles J (1993) The no-self theory: Hume, Buddhism, and personal identity. Philos East West 43:175–200

Gunkel DJ (2012a) "Communication and artificial intelligence: opportunities and challenges for the 21st Century," communication + 1: vol 1, article 1. http://scholarworks.umass.edu/cpo/vol1/iss1/1. Accessed 3 Apr 2017

Gunkel DJ (2012b) The machine question. In: Critical perspectives on AI, robots, and ethics. MIT Press, Cambridge, MA

Hayles K (1999) How we became posthuman: virtual bodies in cybernetics, literature and informatics. University of Chicago Press, Chicago

Heidegger M (1954) The question concerning technology and other essays. 1954. Trans. William Lovitt. Harper Colophon, New York

Heo HH, Kim MS (2013) The effects of multiculturalism and mechanistic disdain for robots in human-to-robot communication scenarios. Interact Stud 14:81–106

Hofstadter D (1980) Gödel, escher, bach: an eternal golden braid. Vintage Books, New York

Hume D (1896) In: LA Selby-Bigge (ed) A treatise of human nature. Clarendon Press, Oxford (**Reprinted from the Original Edition in threevolumes and edited, with an analytical index**)

Humphreys C (1985) Zen: a way of life. Hodder & Stoughton, London

Jastrow R (1983) The enchanted loom: mind in the universe. Simon & Schuster, New York, NY

Kasulis TP (1981) Zen action zen person. University of Hawai'i, Honolulu

Kim MS, Kim EJ (2013) Humanoid robots as "The Cultural Other": are we able to love our creations? AI Soc 28:309–318

Kim MS, Sur J, Gong L (2009) Humans and humanoid social robots in communication. AI Soc 24:317–325

Kim MS, Gong L, Saito N, Nishigaya K, Cabico M, LaFontaine P (2011) The role of self-construal on preferred communication styles with humanoid robots. Int J Humanoid Rob 8:359–374

Kopf G (2001) Beyond personal identity: Dogen, Nishida, and a phenomenology of no-self. Routledge Studies in Asian Religion, New York, NY

Kurzweil R (1999) The age of spiritual machines: when computers exceed human intelligence. Penguin Books, New York

Kurzweil R (2005) The singularity is near: when humans transcend biology. Viking, New York

Lanza R, Berman R. (2010) Biocentrism: how life and consciousness are the keys to understanding the universe. Benbella Books, Dallas

Lombardo T (2017) Future consciousness: the path to purposeful evolution. Changemakers Books, Alresford, UK

Lusthaus D (1989) A philosophic investigation of the Chéng Wei-Shih Lun: Vasubandhu, Hsuan Tsang and the transmission of Vijnapti-Matra (Yogacara) from India to China. Ph.D. Diss., Temple University. University Microfilms International, Ann Arbor, MI

Malik K (2001) What is it to be human?: what science can and cannot tell us. Academy of Ideas, London

Mazlish B (1993) The fourth discontinuity. Yale University Press, New Haven and London

McLuhan M (1964) Understanding media: the extensions of man. McGraw-Hill, New York

Moravec H (1999) Robot: mere machine to transcendent mind. Oxford University Press, New York

Olivier B (2008) When robots would really be human simulacra: love and the ethical in Spielberg's AI and Proyas's I. Robot Film Philos 12:30–44

Pepperell R (2003) The posthuman condition: consciousness beyond the brain. Intellect Books, Bristol

Pepperell R (2005) Posthumans and extended experience. J Evolut Technol 14:43–53

Poole DL, MacWorth AK (2010) Artificial intelligence: foundations of computational agents. Cambridge University Press, Cambridge, UK

Pribram K (2013) The form within. Prospecta Press, Westport, CT

Roszak T (1986) The cult of information. Pantheon Books, New York

Sims CA (2009) The dangers of individualism and the human relationship to technology in Philip K. Dick's *Do Androids Dream of Electric Sheep?*. Sci Fict Stud 36:67–86

Somerville M (2006) The ethical imagination: journeys of the human spirit. House of Anansi, Toronto

Tofts D (2003) Avatars of the tortoise: life, longevity, and simulation. Digit Creat 14:54–63

Varela FJ, Thompson E, Rosch E (1991) The embodied mind. MIT Press, Cambridge

Virillo P (1994) The vision machine, trans. Julie Rose. British Film Institute, London