The Singularity: Beyond Philosophy of Mind

Eric Steinhart

Published as Steinhart, E. (2012) The Singularity: Beyond philosophy of mind. *Journal of Consciousness Studies 19* (7-8), 131-137.

ABSTRACT: Thought about the singularity intersects the philosophy of mind in deep and important ways. However, thought about the singularity also intersects many other areas of philosophy, including the history of philosophy, metaphysics, the philosophy of science, and the philosophy of religion. I point to some of those intersections.

According to Chalmers (2010: 10), thought about the singularity is closely associated with many important philosophical topics. As expected, Chalmers is mainly interested in topics that fall in or near the philosophy of mind (e.g. consciousness, personal persistence). Those topics are of great importance. But the singularity is also connected with topics in the history of philosophy, metaphysics, philosophy of science, and philosophy of religion. And it is connected with topics in ethics and political philosophy. Chalmers is correctly disappointed by the dearth of philosophical interest in the singularity (2010: 9-10). I hope to arouse further philosophical interest by writing in a deliberately provocative way about the links between the singularity and topics outside of the philosophy of mind.

As with everything in the future, the singularity has a past. For the historian of philosophy, the singularity may inspire new ways of looking at some old traditional themes. One such theme is the deep *metaphysics* of intelligence. This metaphysics goes back to the *nous* of Anaxagoras and the *logos* of Heraclitus; it was developed in fascinating ways by the Stoics. Of course, theists have long argued that intelligence plays a foundational role: our universe exists because it was designed by a divine intelligence. And there is a long tradition that portrays our universe as a progressive process in which intelligence is always rising to greater heights (Lovejoy, 1936: ch. IX). For the Hegelian, the singularity looks much like the final self-realization of Spirit in absolute knowing (Zimmerman, 2008). Peirce develops an evolutionary cosmology in which the universe makes endless progress towards a final singularity (Hausman, 1993). Adams used reasoning strikingly like that of Kurzweil to develop his *law of acceleration* (1904, 1909). After World War II, one leading figure in the philosophical history of the singularity is Teilhard de Chardin (1955).

Also after World War II, much work related to the singularity is done in the context of *physical eschatology* (Cirkovic, 2003). Barrow and Tipler sketch a view of the future of the universe as converging towards an ultimate Omega Point (1986). For them, the singularity is the final phase in a long evolutionary process. However, evolution is not random; on the contrary, it is *progressive*. It is governed by some hypothesized *law of optimality*. For Barrow and Tipler, this law is the *Final Anthropic Principle*: "Intelligent information-processing must come into existence in the Universe, and, once it comes into

existence, it will never die out" (Barrow & Tipler, 1986: 23). Tipler later interprets this law as his *Eternal Life Postulate* (1995: 11). For Kurzweil, this law is the *Law of Accelerating Returns:* "As order exponentially increases, time exponentially speeds up" (Kurzweil, 1999: 30; 2005: chs. 1 & 2). Kurzweil thinks of evolution as a self-bootstrapping positive feedback loop in which *value* is always exponentially increasing (1999: 32).

According to their authors, these laws of optimality, like other natural laws, were inferred from empirical data by inductive arguments. For Barrow and Tipler, the induction is by inference to the best explanation. They say that some anthropic principle is the best explanation for the apparent fine-tuning of our universe for life. Ultimately, they derived their Final Anthropic Principle from an atheistic version of the design argument. For Kurzweil, the induction seems to be mostly empirical generalization (1999; 2005). Advocates of the singularity say that these laws can be used to predict the singularity much as the laws of Newtonian gravitation were used by Halley to predict his comet. Kurzweil is well-known for his detailed timelines (1999: 261-280; 2005: chs. 1 & 2). Of course, either the singularity will happen as predicted or it will not. If it does not, skeptics are entitled to reject the alleged laws as mere pseudo-science. However, singularitarians have an intriguing reply. The reply is to push these laws deeper: they should not have been formulated as laws about our universe. Correctly formulated, they work at much greater depths.

Perhaps these laws are related to the laws of optimality discussed by recent philosophers. Leslie argues that ethical requirements have creative power (1970, 1979). Rescher states the law of optimality like this: "whatever possibility is for the best is *ipso facto* the possibility that is actualized"(2000: 815; 2000: 814-821; 1984: ch. 2). At least for *axiarchists* like Leslie and Rescher, the laws of optimality are deeper than any natural laws. They are metaphysical principles that are not empirically falsifiable. Further reflection on the laws of optimality may lead the singularitarian to look at Leibniz's cosmological argument (1697: 84-85). Leibniz says that there must be some ultimate sufficient reason for existence. Of course, Leibniz also says that this reason must be grounded in God. But the singularitarian is more likely to prefer an atheistic account: Why is there something rather than nothing? Because there is some law of optimality that eventually ensures the singularity.

Whether or not the singularity occurs as predicted in our universe, thinking about the singularity leads to new ways for the philosopher of religion to think about the old design arguments for God. Many writers say the design arguments lead to a regression of designers (e.g. Hume, 1779: 72-73; Doore, 1980: 153-155; Dawkins, 2008: 136, 146, 188). Our god designed our universe; but our god was itself designed by some previous god. There are three ways this regression can go: (1) each previous god is more divine than the next god; (2) each previous god is just as divine as the next god; and (3) each previous god is less divine than the next god. Since only the third option avoids some obvious fatal flaws, the regression of gods eventually bottoms out in some least god. When this regression is turned around, it presents itself as a progression of ever more divine gods. This progression is analogous to the progression of ever greater artificial

intellects discussed by singularitarians. It is often said that the singularity will be achieved through the *recursive self-improvement* of artificial intellects (Good, 1965; Kurzweil, 2005: 27-28; Schmidhuber, 2007; Chalmers, 2010: 11-22). Anyone interested in the design argument is free to apply recursive self-improvement to the progression of gods.

This leads to an *evolutionary theogony*. Since evolution is self-bootstrapping, it brings some initial god into being out of nothingness. Adopting what Leslie calls extreme axiarchism (1970: 286; 1979: 6), one might argue that an abstract law of optimality brings this initial god into existence. The initial god exists because it ought to. The initial god has some minimal level of perfection (it is minimally benevolent, intelligent, and powerful). However small this minimal level may be, it is sufficient for the design and creation of the next more divine god. The result is a series of gods in which each previous god designs and creates the next god. The next god is always more divine than the previous god. It is interesting to note that this progression of gods is compatible with Aquinas's Third Way (*Summa Theologica*, Part 1, Q. 2, Art. 3). And yet, surprisingly, this evolutionary theogony is atheistic. Every god is surpassed by some greater god. Just as there is no biggest number, so there is no maximally perfect being – the god of theism does not appear in this sequence. This evolutionary theogony is an atheistic interpretation of the Leibnizian doctrine of the *striving possibles* (Leibniz, 1697; Rescher, 1991: 171-172, 188, 191; Blumenfeld, 1981). And these gods are just big computers. They are not the objects of revealed religion – they are the objects of pure computer science.

The evolutionary theogony entails the production of an endless series of ever better universes. Just as our local god has designed and created our universe, so every god designs and creates its own universe. Just as later gods are more divine, so later universes are better. Philosophers who have reflected on the ontological argument have described such sequences of greater gods making greater universes (Henle, 1961). The classical problem of evil vanishes: every god does the best it can; but no god is unsurpassable. For the singularitarian who is willing to embrace this evolutionary theogony, the singularity becomes much harder to refute. The evolution of intelligence is at work on a super-cosmic scale. Within every next universe, the internal evolutionary process runs farther. Universes can be classified on a Kurzweil scale (2005: 14-21): universes in the n-th rank evolve only through Kurzweil's n-th epoch. Perhaps our universe is just too primitive for the singularity – it merely lies in the fifth rank of universes. Not to worry: the singularity will happen in some near descendent of our universe. It is reasonable to complain that this type of thinking is extremely speculative – but any philosopher has to accept *some* theory of the ultimate origination of concrete actuality. The only question is how well this evolutionary theogony compares with its rivals (which are primarily theistic).

Uploading is often associated with the singularity (Kurzweil, 2005: 198-204; Chalmers, 2010: 41-63). For Chalmers, and for other philosophers of mind, uploading is a lens that focuses thought on the nature of consciousness and problems of personal persistence. But uploading is connected with many topics outside of philosophy of mind. Uploading

raises fascinating issues for political philosophers. If we are going to be uploaded into some cyber-world, surely it will be politically structured. Obviously, speculation about cyber-politics is analogous to classical utopian speculation. Human writers have designed many utopias (e.g. Plato's Republic, Bacon's Bensalem, and so on). But what sorts of utopias will the great artilects design? Chalmers considers a Kantian link between intelligence and benevolence (2010: 36-37). Bostrom suggests that super-human artilects will also be super-human ethical reasoners (2003: 280). On this line of thought, the cyber-world will have heavenly laws. Can we even speculate about those laws? And how would they be enforced? One idea is that the artilects may alter the ethical natures of uploaded humans. Presumably, it will be easy for the artilects to edit our software brains. Perhaps they will get rid of our old sinful natures – editing out the lust and aggression. Perhaps we will also be *morally enhanced* (Douglas, 2008; Faust, 2008). It is easy to see that uploading raises many fascinating ethical and political questions.

Uploading is a secular model of the resurrection theory of John Hick (1976: ch. 15). And many Christian writers have used uploading (or something very much like it) as an analogy for resurrection (Polkinghorne, 1985: 180-181, 2002; Reichenbach, 1978: 27; Mackay, 1997: 248-249). An adventurous thinker might work to link the evolutionary theogony with Hick's *pareschatology* (1976: ch. 21). If the evolutionary theogony is not compatible with orthodox Christian theology, it may at least be compatible with certain radical expressions of Christian soteriology. Many writers associated with the singularity have also explicitly defended resurrection theories. Moravec talks about resurrection (1988: 122-124). Tipler says we will be resurrected in the computers of the far future (1995: chs. 9-13). And Kurzweil says that he will used advanced technology to resurrect his dead father (Kushner, 2009: 61). Since uploading is technically possible, the resurrection of the body is intelligible. Uploading allows anyone interested in bodily life after death to replace a religious mystery with techno-scientific intelligibility. Strikingly, the resurrection of the body becomes a sub-topic in pure (and perhaps applied) computer science.

Philosophers of science have long known that scientific theories often posit ideal objects. Physical theories discuss frictionless planes and ideal gasses. And pure computer science posits ideal computers (e.g. Turing machines). For the modal realist, these ideal objects exist in other possible universes. Although these ideal objects do not exist in our universe, they nevertheless remain legitimate objects of scientific discussion. They are included in the scopes of our best scientific theories. From the point of view of metaphysics and philosophy of religion, whether or not the singularity actually happens in the future of our universe is not very interesting. What is far more interesting is that singularitarian thought suggests that many of the objects and processes that once lay in the domain of revealed religion now lie in the domain of pure computer science.

Notes

References

Adams, H. (1909) The rule of phase applied to history. In H. Adams & B. Adams (1920) *The Degradation of the Democratic Dogma*. New York: Macmillan, 267-311.

Adams, H. (1904) A law of acceleration. In H. Adams (1919) *The Education of Henry Adams*. New York: Houghton Mifflin, ch. 34.

Barrow, J. & Tipler, F. (1986) *The Anthropic Cosmological Principle*. New York: Oxford University Press.

Blumenfeld, D. (1981) Leibniz's theory of the striving possibles. In R. S. Woolhouse (Ed.), *Leibniz: Metaphysics and Philosophy of Science*. New York: Oxford University Press, 77-88.

Bostrom, N. (2003) Ethical issues in advanced artificial intelligence. In S. Schneider (2009) *Science Fiction and Philosophy*. Malden, MA: Wiley-Blackwell, ch. 22.

Chalmers, D. (2010) The Singularity: A philosophical analysis. *Journal of Consciousness Studies* 17:7-65.

Cirkovic. M. (2003) Resource letter: PEs-1: Physical eschatology. *American Journal of Physics 71*(2):122-133.

Dawkins, R. (2008) *The God Delusion*. New York: Houghton-Mifflin.

Doore, G. (1980) The argument from design: Some better reasons for agreeing with Hume. *Religious Studies 16* (2), 145-161.

Douglas, T. (2008) Moral enhancement. *Journal of Applied Philosophy 25* (3), 228-245.

Faust, H. (2008) Should we select for moral enhancement? *Theoretical Medicine and Bioethics 29,* 397-416.

Good, I. (1965) Speculations concerning the first ultraintelligent machine. In Alt, F. & Rubinoff, M. (Eds.) *Advances in Computers, Vol. 6*. New York: Academic Press.

Hausman, C. (1993) *Charles S. Peirce's Evolutionary Philosophy*. New York: Cambridge University Press.

Henle, P. (1961) Uses of the ontological argument. *The Philosophical Review 70* (1), 102-109.

Hick, J. (1976) *Death and Eternal Life*. New York: Harper & Row.

Hume, D. (1779 / 1990) *Dialogues Concerning Natural Religion*. New York: Penguin.

Kurzweil, R. (1999) *The Age of Spiritual Machines*. New York: Penguin.

Kurzweil, R. (2005) *The Singularity is Near: When Humans Transcend Biology*. New York: Viking.

Kushner, D. (2009) When man and machine merge. Interview with Ray Kurzweil. *Rolling Stone* (19 February 2009), 57-61.

Leibniz, G. W. (1697/1988) On the ultimate origination of the universe. In P. Schrecker & A. Schrecker (1988) *Leibniz: Monadology and Other Essays*. New York: Macmillan Publishing.

Leslie, J. (1970) The theory that the world exists because it should. *American Philosophical Quarterly 7* (4), 286-298.

Leslie, J. (1979) *Value and Existence*. Totowa, NJ: Rowman & Littlefield.

Lovejoy, A. (1936) *The Great Chain of Being*. Cambridge, MA: Harvard University Press.

Mackay, D. (1997) Computer software and life after death. In P. Edwards (Ed.) *Immortality*. Amherst, NY: Prometheus Books, 248 – 249.

Moravec, H. (1988), *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press.

Polkinghorne, J. C. (1985) The scientific worldview and a destiny beyond death. In G. MacGregor (Ed.) *Immortality and Human Destiny: A Variety of Views*. New York: Paragon House, 180 - 183.

Polkinghorne, J. C. (2002) Eschatological credibility. In T. Peters, R. J. Russell, and M. Welker, (2002) *Resurrection: Theological and Scientific Assessments*. Grand Rapids, MI: Eerdmans Publishing, 43 – 55.

Reichenbach, B. (1978) Monism and the possibility of life after death. *Religious Studies 14* (1), 27 – 34.

Rescher, N. (1984) *The Riddle of Existence: An Essay in Idealistic Metaphysics*. New York: University Press of America.

Rescher, N. (1991) *G. W. Leibniz's Monadology: An Edition for Students*. Pittsburgh, PA: University of Pittsburgh Press.

Rescher, N. (2000) Optimalism and axiological metaphysics. *The Review of Metaphysics 53* (4), 807-835.

Schmidhuber, J. (2007) Godel machines: Fully self-referential optimal universal self-improvers. In B. Goertzel & C. Pennachin (Eds.) (2007) *Artificial General Intelligence*. Berlin: Springer-Verlag, 199-226.

Teilhard de Chardin, P. (1955 / 2002) *The Phenomenon of Man*. Trans. B. Wall. New York: Harper Collins. Originally written 1938-1940.

Tipler, F. (1995) *The Physics of Immortality: Modern Cosmology, God and the Resurrection of the Dead*. New York: Anchor Books.

Zimmerman, M. (2008) The singularity: A crucial phase in divine self-actualization? *Cosmos and History: The Journal of Natural and Social Philosophy 4* (1-2), 347-370.