An Alternative to Technological Singularity?

George Petelin

Abstract

Technological Singularity, as first articulated by Science Fiction writer Vernor Vinge, and predicted by futurist Ray Kurzweil to occur in just fourteen years from now, has provided lurid scenarios for the post-cyberpunk genre. Writers such as Charles Stross and Neal Stephenson have described worlds where (post-) humans are entirely dependent on machines that have developed beyond their comprehension and have achieved an uncanny verisimilitude to sentience.

While much scepticism prevails in the academic community, Kurzwiel's portrayal of the coming phenomenon and that of the SF fraternity is predicated on a faith in progress. Even though, unlike Kurzweil, the science fiction writers invariably find dystopian characteristics in this future, both envisage an unbroken continuity of technological development. And both dream of machines achieving a level of sophistication beyond human comprehension or of humans merging with their creations to the point of being essentially indistinguishable from them.

However, I will argue that both over-enthusiasm and scepticism regarding this portrayal of robot entities smarter than humans impede attention to a very real and mundane threatening phenomenon: the accelerating complexity of technologies in general and the increasing inability of human institutions to apply these technologies critically, ethically, and advantageously. Technology does not have to be 'smarter' than us to become incomprehensible, only more complex and change at a rate faster than social processes can match. This point, I will argue, we have already reached but refuse to admit.

But there may yet be another turning-point at which we decide that the rate of change needs to decelerate; our mechanisms are actually much 'dumber' that we ourselves could be if only we used them judiciously and stopped allowing market-driven hyperbole to persuade us they should be adopted indiscriminately. Unfortunately, such a 'singularity' would need a redefinition of both progress and machine intelligence and will most likely be precipitated by a catastrophe, a cataclysmic failure of the ultimately 'dumb' technologies on which we have been persuaded to be utterly dependent.

Key Words: Cyberculture, progress, technological singularity, utopia, dystopia.

*****

1. Singularity

The protagonists in Vernor Vinge's 1986 novel *Marooned in Realtime,* attempting to explain the disappearance of a preceding super civilization, sum up both the promise and the apprehension that the singularity evokes:

> By 2200, we could increase human intelligence itself. And intelligence is the basis of all progress. My guess is that by midcentury, any goal—any goal you can state objectively, without internal contradictions, could be achieved. And what would things be like fifty years after that. There would still be goals and there would still be striving, but not what we could understand.
>
> To call that time `the Extinction' is absurd. It was a Singularity, a place where extrapolation breaks down and new models must be applied. And those new models are beyond our intelligence.[1]

In Vinge's novel, civilizations that attain singularity vanish from the universe, at least to the perception of everyone else. This is because their intelligence is so far above the ordinary that their civilization is beyond our comprehension and maybe also because they have chosen a totally different plane of existence. The catch is that, by definition, the nature of this state is even beyond theorisation. And the second catch is that, because this advance is the result of exponentially accelerating returns, it will take us almost completely by surprise when it reaches a point on the graph where its rise becomes vertical—virtually of infinite speed. Is this threshold then desirable?

Vinge is a scientist as well as a writer and the term he first coined in fiction has come to be taken so seriously among technophiles as to generate somewhat of a cult and a range of interpretations and extrapolations among professional futurists.

What gets debated is the *type* of intelligence that could come about and whether it will remain 'human' in the sense we now understand ourselves to be. Will the technology advance with or without its creators? One scenario could be that we will simply be able to hasten biological processes that improve health, increase longevity, and ensure enough increased reasoning power to maintain some control over technology while still remaining essentially organic creatures. These creatures, however, because of genetic manipulation, might be far from being human as we now know it. Another prognosis for the Information Technology revolution suggests that 'humans' of the future may be extensively augmented by non-organic hardware, or even populated by sub-microscopic nano-computers. A combination of these scenarios is deemed highly likely by futurists dubbed Singularians because of the imminent convergence of biotechnology and digital computing. But what is also proposed is that humans may not be able to keep up with their machines and be replaced by them altogether. To do this, machines would need something equivalent to sentience, a consciousness, a sense of personal purpose apart from that of their creators, and to take control of their own reproduction and evolution—improve themselves rapidly in ways unimagined by their original creators. In other words, robots would have to invent a form of sex and perhaps of 'robot-eugenics'. Yet another hypothesis is that human

intelligence could merge with the big global machine, be uploaded, to frolic forever in a virtual environment. But most intriguing is the notion that the exponentially growing internet might reach a point where its complexity generates an 'emergence' of a conscious technological entity.

I examine some criticisms of these predictions, but what I mainly want to argue is that it distracts from more immediate problems: the fact that an increasing complexity and convergence of technologies incontrovertibly exists and precipitates issues quite apart from the promise or threat of its own sentience.

The concept of a singularity based on machine super intelligence was first proposed by Vinge in a 1981 novella titled 'True Names'[2] and was immediately endorsed by Marvin Minsky a cognitive scientist working on Artificial Intelligence.[3] In 1993 Vinge set out its principles within a provocative paper titled *The Coming Technological Singularity*. [4] As well as Vinge, several futurists stand out as proponents of singularity the chief which is Ray Kurzweil, author of books such as *The Age of Spiritual Machines*[5], and *The Singularity Is near: When Humans Transcend Biology* [6] and articles such as *The Law of Accelerating Returns*[7]. Other such writers include robotics futurist Hans Moravec[8], popular science and SF writer Damien Broderick,[9] and arch techno-optimist and proponent of media convergence James Canton.[10] All of these predict, as does Vinge[11], that the singularity is imminent within the next two decades.

Science Fiction writers have pounced on the concept of Singularity with glee but invariably also with criticism: in the cyberpunk era they produced accounts of dystopian societies approaching, or left behind by, singularity and, in the post-cyberpunk phase, joyfully satirised the predictions and possible consequences of Singularity.[12]

2. The Terms of Debate and Why Machines are 'Dumb'

Although there had long been little academic attention to this phenomenon, serious interdisciplinary discussion is beginning to take place. However, as among the professional futurists, technological over-enthusiasm continues to limit criticism to speculation instead of addressing what is actually occurring. I use the following debate started by philosopher David Chalmers to demonstrate the explicit and implicit issues at stake.

Setting out to critique the dominant Singularian paradigm, Chalmers starts to set out its parameters by quoting I.J. Good in his 1965 article 'Speculations Concerning the First Ultraintelligent Machine':

> Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultra- intelligent machine could design even better machines; there would then unquestionably be an 'intelligence explosion', and the intelligence of man would be left far behind. Thus the first ultraintelligent machine is the last invention that man need ever make. [13]

Accepting this definition as his starting point, Chalmers acknowledges, but bypasses, the question of what is intelligence and instead interrogates the timeframe promised by Singularians. Rejecting hardware extrapolations based on 'Moore's law', he writes 'My own view is that the history of artificial intelligence suggests that the biggest bottleneck on the path to AI is software, not hardware: we have to find the right algorithms, and no-one has come close to finding them yet'.[14]

In brief, Chalmers does not reject the possibility of AI and even super-AI since he holds that the human is a biological machine and therefore there is no reason why we cannot duplicate what evolution has achieved. He also considers it a likelihood that recursive machines will lead to an explosion of intelligence and hence to greater happiness. This is despite the fact that the progress of science has not thus far had a good track record in this last regard, when all is considered.

Now, algorithms are precise steps or instructions, principles of operation that allow calculation to take place. They are unlike heuristics which are the flexible rules of thumb that guide embodied human interaction within the unpredictable totality of their complex sensory and social environment. Rather than mathematical operations heuristics are more in the nature of personal disciplines.

A kind of oxymoron, the term 'heuristic algorithm', has been coined by computer scientists to describe programs that that employ random choices and fuzzy logic because it is theorised that they can arrive at a near optimum solution quickly, although a perfect solution only in infinite time. Like heuristics employed by humans, they are thought to be the path to new discoveries and inventions. Computer heuristic algorithms nonetheless remain strictly defined calculations performed on specified forms of data and not interpretations or judgements based on a totality of experience. Moreover, the criteria for determining whether their solution is 'optimum' are necessarily finite and limited or else one would still have to wait for an infinite time to make that judgement. In an interview regarding the possibility of AI, Noam Chomsky, whose theory of the hierarchy of formal grammars[15] is integral to distinguishing simple from complex algorithms, says:

> What's a program? A program is a theory; it's a theory written in an arcane, complex notation designed to be executed by the machine. What about the program, you ask? The same questions you ask about any other theory: Does it give insight and understanding? These theories don't.[16]

Observing that, in addition to data processing, computation invariably requires humans to interpret 'multitudinous', 'vague', and 'ambiguous' real data and circumstances in terms of propositions which then need reduction to their 'elementary denials' before processing, and ultimately human interpretation of results, logician John Venn in 1881 dismissed the possibility that 'any contrivances at present known or *likely to be discovered* [my italics] really deserve the name of logical machines'.[17] And virtually every definition of an algorithm stresses its precision at the expense of ambiguity and subtlety that characterises human thinking and experience. Daniel Dennett sums up the limitation of algorithms:

> No matter how impressive the products of an algorithm, the underlying process always consists of nothing but a set of individually mindless steps succeeding each other without the help of any intelligent supervision; they are 'automatic' by definition: the workings of an automaton.[18]

As John Searle points out, in aspiring to intelligence, computers attempt to deal with symbols but symbols in human society derive their meaning not only from the grammar of their language but out of the circumstances of their performance and reception. 'Nothing is intrinsically computational' says Searle, 'Computation exists only relative to some agent or observer who imposes a computational interpretation on some phenomenon …'.[19]

The equivocation that confuses algorithmic calculation with intelligence leads to both an anthropomorphism and an overreliance on technology. Chalmers proceeds to hypothesise how the threat of machines may be controlled by designing them to have 'values' or limits in their functionality.

And how does he believe it should be determined what values machines should have? Why, by modelling their likely progress in a computer! Again and again, a technological solution, adjusting the algorithm in the machine, is called to the rescue rather than an interrogation of how *we* use machines and how our political economy encourages us to use them.

Taking a somewhat abrupt leap from computational 'intelligence' to consciousness, Chalmers sees it as inevitable that the conscious super-AI will successfully negotiate with humans to free itself from whatever constraints it regards as irrational. The only way to reintegrate ourselves into this regime, according to Chalmers, would be to upload ourselves into the machines. He then proceeds to discuss the merits of various ways of 'uploading' a human consciousness into the computer: whether this would result in immortality or duplication of living beings and related questions.

Much of the debate that occurs among Singularists is based on circular arguments. Computation is considered to be intelligence as a first premise. Intellect, consciousness, and humanity, then all become uncritically conflated with this 'intelligence'. For example, Jürgen Schmidhuber, in replying to Chalmers's paper, categorically asserts that

> The scientific way of measuring intelligence involves measuring problem solving capacity. There are mathematically sound ways of doing this, using basic concepts of theoretical computer science, all of them avoiding the subjectivity of the ancient and popular but scientifically not very useful Turing test, which essentially says 'intelligent is what I feel is intelligent.[20]

Schmidhuber's notion of intelligence derives from computer science rather than from the humanities, which surely have a primary stake in its definition. Intelligence, in this context, is entirely a matter of calculation rather than phenomenological experience. If computers begin as the yardstick of intelligence, they will surely also end up being judged the epitome of its attainment. Thus, almost predictably, Schmidhuber concludes that 'there already exists the blueprint of a Universal AI which will solve almost all problems almost as quickly as if it already knew the best (unknown) algorithm for solving them'.[21]

Conflating hardware with software and neurological structure with intellect, Schmidhuber argues that such a Universal AI would 'create feature hierarchies, lower level neurons corresponding to simple feature detectors similar to those found in human brains, higher layer neurons typically corresponding to more abstract features, but fine-grained where necessary'.[22]

Another reply to Chalmers, by Francis Heylighen, claims to adopt a 'situated and embodied view of cognition', noting that 'intelligence is rooted in a high-bandwidth, sensory-motor interaction with the outside world'.[23] Pointing out Chalmers's reductionist neglect of holistic embodied experience and accusing him and other Singularity theorists of mind-matter dualism, Heylighen nonetheless also proposes a technological solution—computer networking.

In networked technologies, he argues,

> the mind is no longer localized in any particular component, but distributed over a massive number of internal and external components which all cooperate in a self-organizing manner. Intelligence (and with it consciousness) can then be seen as an emergent phenomenon of coordination between these processes, an integrated manner of dealing with an enormous amount of bits and pieces that together determine an individual's experience of its situation, and that define a potential problem to be dealt with.[24]

As a result of this consideration, Heylighen notes that

> In the last decade or so, the focus of AI has therefore shifted towards machine learning and data mining, i.e. letting the computer program itself extract knowledge from the huge amounts of data available in specialized databases and on the web.[25]

Acknowledging that most data is entered by humans, Heylighen sees the AI's lack of a biological body only a temporary handicap. Apart from making the absurd conclusion that sheer complexity inevitably leads to the emergence of consciousness (we can ask for example why the brain of a corpse, no less complex than when it was living, lacks consciousness), he ignores a multitude of sinister factors implicit in the accomplishment of this strategy: the ethical issue of mass surveillance, the fact that increasing obligations to enter data reduce us to disempowered servants of the techno-economic complex, the fact that much of this data is reductivist, distorted, or sheer misinformation, and the unwarranted assumption that the data will be used for benevolent problem-solving. Nonetheless, Heylighen judges that

> …there are plenty of observations as well as theoretical arguments for believing that this collective system, which may be called the Global Brain, is not only intelligent, but becoming quickly more intelligent. The reason is that its

> self-organization is facilitated and accelerated by the seemingly unpreventable processes of globalization and of the increasing spread of information and communication technology. The result of these processes is that information about what is happening on this planet becomes ever more easily available, helping us to make better, more informed decisions, and to tackle more complex problems.[26]

Not only is technology thus rendered unproblematic but, together with it, the whole process of globalisation.

3. What are the real threats?

The entire debate is confused through an inability to define or admit what is *human* intelligence, what are intellectual activities, and what constitutes being 'better' at them. It is repeatedly implied that intellect can be reduced to mere calculation. And this sort of intellect is independent of social and personal parameters. Yet human achievements in our society by definition are measured in terms of the quality of human satisfaction that they generate in relation to a specific individual or a group within society, rather than in terms of some complexity or quantity of problem solving. Some extremely simple inventions and works of art requiring little calculation are considered achievements of genius. Despite this, and the essentially dispassionate nature of the algorithms on which 'robots' of the future must be based, they are invariably conceived as some kind of homunculi, with potential personal and social goals and desires equivalent to the most superficial aspirations attributed to humans in western society.

Although they may mimic human behaviour, there can be no spontaneous affect, no unconscious nor philosophical dimension to these calculating machines. There are no consequences that they can be committed to except abstractly 'better' performance according to ethically and politically unspecified criteria. They are the ultimate triumph of instrumental reason without purpose. They are, in short, managerial bureaucrats, 'objectively' carrying out tasks in zombie fashion without any commitment to intrinsic experience. The frightening thing is that we are required to be increasingly like them. Machines are not turning themselves into people, it is we who are turning into mindless performers of algorithms in the service of blind accumulation and growth.

I would argue that the process of turning us into robots has already begun, but that it is not of the same nature and as inevitable or desirable as argued by Morovac.[27] What is crucial is not that our bodily parts, or even our calculating ability, are being replaced, but that we are abrogating our judgement and our human rights to the idol of technology and its commercial interests. Thus any concern about the potential consciousness and tyranny over humans of robots is sadly misplaced. The real issue is the rights of human individuals and communities in relation to commercial interests vested in algorithm-operated corporations, accounting systems, software services, and the like. It is precisely their lack of consciousness that makes these mechanisms a tyrannical threat to humanity.

4. Techno-opotimism and the inevitable collapse of global IT

The adoption of extreme optimistic positions regarding technology by industry figures such as Kurzweil and Minsky can be explained simply. Robert Geraci in his research for the book *Apocalyptic AI*, found that researchers in robotics were concerned with far more mundane issues of simply getting a robot to work and in the main were not anxious about their transcendental or ethical consequences.

> Eventually, I gained a clearer understanding of how prestige and public approval of robotics/AI research plays a role in Apocalyptic AI—robotics and AI enjoy government support as a consequence of the fantastic promises made by Apocalyptic AI authors. Promises of intelligent robots and uploaded consciousness could have replicated successfully through science fiction without ever mixing so closely with laboratory science, as they do in Apocalyptic AI pop science books. The value of the apocalyptic imagination lies in its power to create excitement in the lay public and government funding agencies. Pop science in general, and Apocalyptic AI in particular, is a—sometimes conscious, sometimes unconscious—strategy for the acquisition of cultural prestige, especially as such prestige is measured in financial support.[28]

There are unfortunate potential consequences of technology convergence, the cult of optimism surrounding it, and its real or apparent exponential progress that have gained little purchase in the debates. One is that exponential growth is not infinitely sustainable. Another, is that when it even falters, and more so if it collapses, everything we do will be affected. However, the hype that attends technology continues to urge us to trust in mindless algorithmic progress and put all our resources into one ultimately fragile basket.

Notes

[1] Vinge, *Marooned in Realtime*.
[2] Vinge, 'True Names'.
[3] Minsky, *The Society of Mind*. Minsky first wrote part of this as an epilogue for Vernor Vinge's 'True Names'.
[4] Vinge, 'The Coming Technological Singularity'.
[5] Kurzweil, *The Age of Spiritual Machines*.
[6] Kurzweil, *The Singularity Is near*.
[7] Kurzweil, 'The Law of Accelerating Returns.'

[8] Moravec, *Robot*.
[9] Broderick, *The Spike*.
[10] Canton, *The Extreme Future*.
[11] Vinge, 'The Coming Technological Singularity.'
[12] Raulerson, 'Singularities.'
[13] Chalmers, 'The Singularity,' 7.
[14] Ibid., 12.
[15] Chomsky, 'Three Models for the Description of Language.'
[16] 'Noam Chomsky on Singularity 1 on 1: The Singularity Is Science Fiction!'
[17] Venn, *Symbolic Logic*, 161.
[18] Dennett, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, 59.
[19] Searle, *Consciousness and Language*, 17.
[20] Schmidhuber, 'Philosophers & Futurists, Catch up,' 175.
[21] Ibid., 177.
[22] Ibid., 180.
[23] Heylighen, 'A Brain in a Vat Cannot Break out,' 126.
[24] Ibid., 130.
[25] Ibid., 132.
[26] Ibid., 138–139.
[27] Geraci, *Apocalyptic AI*, 30.
[28] Geraci, *Apocalyptic AI*, 3.

Bibliography

Broderick, Damien. *The Spike: How Our Lives Are Being Transformed by Rapidly Advancing Technologies*. 1st ed. New York: Forge, 2001.
Canton, James. *The Extreme Future: The Top Trends That Will Reshape the World for the next 5, 10, and 20 Years*. New York: Dutton, 2006.
Chalmers, David. 'The Singularity: A Philosophical Analysis.' *Journal of Consciousness Studies* 17, no. 9–10 (2010): 7–65.
Chomsky, Noam. 'Three Models for the Description of Language.' *Information Theory, IRE Transactions on* 2, no. 3 (1956): 113–24.
Dennett, Daniel C. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. London: Penguin Books, 1996.
Geraci, Robert M. *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. New York: Oxford University Press, 2010.
Heylighen, Francis. 'A Brain in a Vat Cannot Break out: Why the Singularity Must Be Extended, Embedded and Embodied.' *Journal of Consciousness Studies* 19, no. 1–2 (2012): 126–42.
Kaczynski, Theodore John. *The Unabomber's Manifesto*. Findlay Publishing Company, 2006.
Kurzweil, Ray. *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*. New York: Viking, 1999.
———. 'The Law of Accelerating Returns.' In *Alan Turing: Life and Legacy of a Great Thinker*, 381–416. Springer, 2004.
———. *The Singularity Is near: When Humans Transcend Biology*. New York: Viking, 2005.
Minsky, Marvin Lee. *The Society of Mind*. New York: Simon and Schuster, 1986.
Moravec, Hans P. *Robot: Mere Machine to Transcendent Mind*. Oxford University Press, 2000.
'Noam Chomsky on Singularity 1 on 1: The Singularity Is Science Fiction!' Accessed February 14, 2014. http://www.singularityweblog.com/noam-chomsky-the-singularity-is-science-fiction/.
Raulerson, Joshua Thomas. 'Singularities: Technoculture, Transhumanism, and Science Fiction in the 21st Century,' 2010. http://ir.uiowa.edu/etd/2968/.
Schmidhuber, Jürgen. 'Philosophers & Futurists, Catch up.' *Journal of Consciousness Studies* 19, no. 1–2 (2012): 173–82.
Searle, John R. *Consciousness and Language*. New York: Cambridge University Press, 2002.
Venn, John. *Symbolic Logic*. London: Macmillan, 1881. https://archive.org/details/symboliclogic00venniala.
Vinge, Vernor. *Marooned in Realtime*. New York: Baen Books, 1987.
———. 'The Coming Technological Singularity.' *Whole Earth Review* 81 (1993): 88–95.
———. 'True Names.' In *Binary Star No. 5*, edited by James R Frenkel. New York: Dell, 1981.