



6. Singularity Contour

By Yehezkel Dror

Contents

1. Preamble	54
2. Not a <i>Technological Singularity</i>	54
3. The Real Singularity.....	57
4. To Be, What to Be, Not to Be	58
5. Fateful Fuzzy Gambles	59
6. Decision Criteria.....	60
7. Global Future-Shaping Super-Elite	62
8. Are Politicians As Such Legitimate Singularity-Shaping Decision-Makers? .	65
9. Required Decision Virtues.....	66
10. Enforcement Regimes	67
11. Global Crises As Indispensable Opportunities	69
12. What Is to be Done Now	70

1. Preamble

- 1.1 This is a first draft of a paper on the main features of the emerging phase leap of the human species, increasingly called the *Singularity*. This special issue of the *Journal of Space Philosophy* provides a welcome opportunity to present it to a distinguished audience, together with a closely related paper on the future Space Epoch by Dr. Bob Krone.
- 1.2 While I endorse and use in this paper the term “Singularity,” personally, I prefer a more professional term that better expresses the contents of the emerging human phase leap, namely *Anthroporegenesis*, as explained in Section 3.2. This is all the more necessary because the term Singularity and its associates, such as *Transhumanity* and *Posthumanity*, are increasingly becoming a flag of convenience for baseless speculations. But there is also a growing body of serious work using these terms (as illustrated in my companion essay “Becoming a Singularity Policy Scientist”), so I adopt them subject to this warning.
- 1.3 Though rudimentary, lacking elaboration, and without academic references, this draft presents the main features, issues, problems, options, and choices related to the Singularity hypothesis, as conceptualized by me. It can also be regarded as an outline exploration of a radically novel and, indeed, revolutionary paradigm presenting the new epoch into which the human species is cascading, however unprepared. But I prefer the less presumptuous term *contour* for exploring what is a conjecture, however well based on the evaluation of ongoing processes.
- 1.4 This draft engages in quite some iconoclasm – casting doubt on widely accepted notions and emphasizing serious lacunae in current discourse on the Singularity. I also transgress against political correctness and deviate from common sense and consensual values. The reason is simple, but it has far-reaching implications:

**“Thinking as usual” while undergoing a metamorphosis is a widely used
speedway leading humankind to avoidable catastrophes.**

- 1.5 All the more so, I need the help of interested readers, who are invited to send their comments and suggestions to the author, at my email address yehezkel.dror@mail.huji.ac.il. To help you to do so on the basis of following my reasoning critically, I put some key statements in emphasized boxes, as illustrated above.

2. Not a *Technological Singularity*

- 2.1 As best presented by Ray Kurzweil in his books, the term Singularity in the present context (as distinct from mathematics and astrophysics) refers to the scientific and technological quantum leap that pushes the human

species into a radically novel mode of being. But what is usually meant, and sometime explicated, is a technological Singularity.

2.2 Without going into the technological specifics, however fascinating, such as the different possible substrata of superintelligence, the following partly overlapping emerging technologies constitute the main relatively realistic dimensions of the (technological) Singularity (leaving for the long-term future esoteric possibilities such as downloading human minds into computers and thus making them nearly eternal):

2.2.1 Artificial general intelligence leading to intelligent and super-intelligent robots and perhaps spiritual machines, increasing or surpassing the mental abilities of human beings, however enhanced. Included are the possibility of human-machine combinations, and also an escalating chain of intelligent robots designing and producing super-intelligent ones, which in turn design and produce super-super-intelligent robots, and so on into an inconceivable future with radical implications for human beings, such as cohabitation, being marginally tolerated, and elimination.

2.2.2 Nanotechnologies, enabling production of nanomachines, including nanorobots that can act within human bodies, prolonging life and enhancing various abilities, or acting as invincible mass killing machines; and surpassing the ambitions of alchemists by transmuting materials and thus eliminating scarcities and rehabilitating the environment, but also creating havoc, such as by reducing the value of gold to that of lead that is cheaply transformable into pure gold.

2.2.3 Genetic engineering enhancing human bodies and minds, prolonging life expectancy, enabling human cloning, making it easy to transmute and synthesize viruses, and much more.

2.2.4 Human-machine interfaces and combinations multiplying human abilities while compensating for bodily and mental deficiencies, up to transforming humans into cyborgs.

2.2.5 Cheap and non-polluting energy that can be easily stored, changing totally all energy-based technologies and tools.

2.2.6 Accelerated and large-scale space exploration, thanks to enhanced human bodies, intelligent robots, and new space traveling technologies (leaving aside long-term possibilities to reach exoplanets and perhaps to expand humanity beyond the Solar System).

- 2.3 There are different opinions on the likely timeline of progress in developing the various Singularity technologies, ranging between a couple of decades and one or two centuries. I personally think that the more critical and radical Singularity innovations require scientific knowledge and technologies far beyond our present reach. Thus, to achieve superintelligence, its currently unknown nature has first to be clarified.
- 2.4 An essential step for doing so is to unravel the riddles of geniuses, which may take a long time if at all possible for human minds. Let me take as an example the Indian mathematician Srinivasa Ramanujan. He was a super-genius, as well put in a biography:

a man who grew up praying to stone deities; who for most of his life took counsel from a family goddess, declaring it was she to whom his mathematical insights were owed; whose theorems would, at intellectually backbreaking cost, be proved true—yet leave mathematicians baffled that anyone could divine them in the first place.¹

No advances in neurosciences or studies of the mind provide plausible explanations of the nature and organic bases of such a genius, as may be essential for designing superintelligent robots.

- 2.5 And what about the four ways to truth, stipulated by Alain Badiou, including art and love (in addition to science and politics in a special sense)? The very term *intelligence* in the term *superintelligence* raises serious questions, and it may be far too narrow, all the more so as neither neurologically nor philosophically are art, and even less so love (as correlated with sexuality and thus both embodied and somehow above the body) included in intelligence.
- 2.6 Therefore, I recommend a good measure of skepticism on predictions that most of the Singularity is around the corner.

My assessment is that it will take at least one or two centuries before salient Singularity technologies become mature, even if not slowed down by civilizational catastrophes or on purpose. And full-scale superintelligence may take even longer, if at all achievable.

- 2.7 Nevertheless, it is very likely that by the end of the 21st century some of the technologies will be mature enough to significantly impact on humanity, providing many blessings accompanied by explosive disruptions and harsh moral and political quandaries.

¹ Robert Kanigel, *The Man Who Knew Infinity* (London: Little, Brown, 1991), Kindle edition, 4.

- 2.8 Thus, within this century, artificial generally intelligent (but not superintelligent) robots are likely to reduce radically labor done by humans, creating mass unemployment. Human enhancement is likely to result in harsh biological inequality between the few who can pay for expensive enhancements, such as significant increases in life expectancy, and the many who will not be able to benefit from such technologies before they become affordable by all – which at best will take a long time. And berserk fanatics are likely to produce mass killing viruses and to use them for extortion and genocide.
- 2.9 The contrasting uses of emerging Singularity technologies for better and worse, as differently defined by the beliefs of diverse parts of humanity, will pose harsh choices requiring unprecedented measures. While most existential risks associated with the Singularity are widely recognized, and countermeasures are analyzed at various academic and policy units, in my view, much more is needed to reduce serious risks to humanity. And achieving the benefit of the Singularity may not be much easier.

3. The Real Singularity

- 3.1 Technology is not an agency, being rather a set of tools based largely on science and hands-on creativity. Technology produces the emerging Singularity, is its landmark, and instantiates it. But the real ontology of the Singularity is different:

The real Singularity is the growing ability of the human species to shape its future evolution as a species, the salient features of the evolution of the biosphere of earth, and its physical surface characteristics.

- 3.2 The human species has influenced its evolution and that of some animal and plant life since its beginnings. Hunting, fire, mating patterns, selective breeding, seed selection, medical knowledge, biopolitics, changing eating patterns – these and related behavior have increasingly impacted on the dynamics of human evolution and parts of its environment, all the more so after the industrial revolution, as recognized by the novel term for our epoch “Anthropocene.” It includes nuclear fusion, the first steps into space, and human-caused global climate changes that constitute the dawn of the Singularity, the core of which is what I call Anthroporegenesis, in the sense of the human species acquiring the technologies that enable it to bring about a new genesis, transforming the human species and its living world radically.

3.3 To put the core ontology of the Singularity clearly:

Humanity as a composite agency is taking charge of critical features of its future evolution as a species, partly displacing natural evolution. It does so thanks to tools provided by leaping science and technology – but the decisions on using these tools are made by humankind.

3.4 The future of our species will continue to depend at least in part on natural events, such as objects from outer space hitting Earth and mega-volcanoes. But the human species is developing technologies it can use to change its biology and minds, to reshape Earth radically, perhaps to settle other planets, and also to terminate the existence of the human species with or without other humanity-generated forms of life taking over.

3.5 Let me recapitulate the critical reformulation of the nature of the emerging Singularity, which is not recognized in current uses of the term but is critical for handling the Singularity, to reduce negative effects and to increase positive ones:

The Singularity is not constituted by technological evolution partly displacing natural evolution, even if phenomenologically this is happening. It is the human species, which acquires the knowledge and tools increasingly enabling it to impact on its future evolution, using the emerging technologies as it chooses. This is not a deterministic process, but one shaped by human discretion. Accordingly, humankind as a composite agency is morally responsible for the Singularity and its consequences, not the scientists and technologists who provide the Singularity tools.

4. To Be, What to Be, Not to Be

4.1 Pondering Singularity scenarios and what to do about them requires a total shift in human perspective:

We have primarily to think, invent, and act in terms of evolutionary processes shaping the future of humanity and the increasing role of human choices in steering them. Therefore, while the welfare of humans now and in the near future continues to be very important, assuring a long-term future for the human species has top priority (unless future generations decide differently, such as letting a super-superior species take over).

4.2 Parts of humanity have faced many critical choices in the past, whether they recognized them as such or not. These have ranged from individual existential choices to collective ones on regimes, economic systems, moral norms, wars or peace, and so on. But few such choices shaped the deeper levels of our long-term history, and none impacted significantly on our nature and evolution as a biological species. All this is changing with the emerging Singularity.

- 4.3 For the first time in the history of life on Earth, a species has the capacity to shape deliberately its future evolution and, consequently, it has to make decisions, explicitly or by default, on the meta-Hamletian question “to be, what to be, or not to be.” And, on a higher order level, humanity has to decide, explicitly or by default, who should make such decisions and shape significantly its evolutionary future, and how to implement such choices effectively.
- 4.4 Some of these issues receive attention, such as discourse on avoiding catastrophes. Also, somewhat elaborated are decision criteria. But most emerging decision forks, including critical ones, are ignored or badly considered – including the meta-issues of who should decide on interventions with human evolution and how to implement such decisions.

5. Fateful Fuzzy Gambles

- 5.1 Historic processes are by their very nature dynamic, non-linear mixtures between necessity and chance. This is true, in various proportions, from the macro-level of cosmic processes to the nano-level of atomic and subatomic events; and also, with adjustments taking into account the important role of “choice,” the life history of individuals. However, totally novel is the increasing importance of human choice in shaping the processes determining the future of the human species:

The future of the human species is increasingly shaped by human choice interacting with necessity and chance. Therefore, application of existential philosophy to humanity as an agency and developing the nature of humankind as a collective deliberative agency are at the core of upgrading human impacts on the future of humanity as a species.

- 5.2 This sounds great, but it may be catastrophic unless human future-impacting choices are of optimal quality. However even optimal choices have results that depend in part on necessity and chance beyond human control. This is all the more so the case in the face of prevailing deep uncertainty. Therefore:

Even optimal human future-shaping choices are inherently and unavoidably “fuzzy gambles, often for high and also fateful stakes”—because the future is in deep uncertainty, reaching wild uncertainty and also inconceivability. This is increasingly the case as we move into the Singularity with its unprecedented phase leaps into the largely unknowable.

- 5.3 If this is the nature also of optimal choices, then all the more so suboptimal choices carry multiplying risks. Taking into account the usually low quality of human choice on complex quandaries, it is far from assured that the increasing power of humanity to shape its future will work for the better. It is

no less likely (to put it relatively optimistically) to result in catastrophes and even collective unintended species suicide.

5.4 Please do not be misled by the fact that humanity is today overall better off by material criteria than ever before. This is true, largely thanks to evolving science and technology, which till recently was in the main very beneficial for humanity without catastrophic risks for the future of the species. Also beneficial have been some governmental policies and social self-regulating processes such as relatively autonomous global markets and mild regimes, such as democracy. But if we take a close look at the actual choices of governments on global issues rather than pious declarations, such as on climate change, then the nakedness of the princes in charge of increasingly critical and perhaps fatal choices is fully revealed.

5.5 A tentative, frightening conclusion is unavoidable:

Choices significantly impacting on the future of humankind suffer from a growing and increasingly dangerous hiatus between the growing magnitude of impacts and the low quality of fuzzy gambling choices by main global decision-makers. This quality deficit is sure to produce global catastrophes. Therefore, radical improvement of critical choices is imperative, together with awareness of unavoidable risks stemming from more powerful technologies, however beneficial they may be if well used.

5.6 The assessment above is at the core of this paper. Therefore, it is explored further in the following sections. Fuller though still partial treatment is provided in other writings by the author.²

6. Decision Criteria

6.1 Given the insights suggested above, it is necessary to consider decision criteria fitting important choices posed by advancing towards the Singularity. Widely accepted, at least verbally, is the precautionary principle, formulated with variations more or less as follows:

The precautionary principle: If the consequences of an action are unknown but judged by some scientists to have even a small risk of being profoundly negative, it's better to not carry out the action than to risk negative consequences.

6.2 This decision criterion follows the minimax game theoretical rule, aiming at minimizing the possible loss for a worst case. Thus, it does not balance potential risks and benefits. Also, it is very vague on the number and quality of pessimistic scientists who have a veto on novel technologies. Little wonder that the precautionary principle is usually not acted upon; and the

² See my short book *For Rulers: Priming Political Leaders for Saving Humanity from Itself* (Washington, DC: Westphalia Press, 2017).

few cases when it was applied, such as the exclusion of mutated seeds by the European Union, are in error, even if they are supported by true Green believers.

6.3 Max More articulates the limitations of the precautionary principle and advocates replacing it with what he calls the *proactionary principle*, which involves balancing the risks of action and inaction. In principle, this is correct. Let me therefore quote a summary of the proactive principle as proposed by him:

1. People's freedom to innovate technologically is valuable to humanity. The burden of proof therefore belongs to those who propose restrictive measures. All proposed measures should be closely scrutinized.
2. Evaluate risk according to available science, not popular perception, and allow for common reasoning biases.
3. Give precedence to ameliorating known and proven threats to human health and environmental quality over acting against hypothetical risks.
4. Treat technological risks on the same basis as natural risks; avoid underweighting natural risks and overweighting human-technological risks. Fully account for the benefits of technological advances.
5. Estimate the lost opportunities of abandoning a technology, and take into account the costs and risks of substituting other credible options, carefully considering widely distributed effects and follow-on effects.
6. Consider restrictive measures only if the potential impact of an activity has both significant probability and severity. In such cases, if the activity also generates benefits, discount the impacts according to the feasibility of adapting to the adverse effects. If measures to limit technological advance appear justified, ensure that the extent of those measures is proportionate to the extent of the probable effects.
7. When choosing among measures to restrict technological innovation, prioritize decision criteria as follows: give priority to risks to human and other intelligent life over risks to other species; give non-lethal threats to human health priority over threats limited to the environment (within reasonable limits); give priority to immediate threats over distant threats; prefer the measure with the highest expectation value by giving priority to more certain over less certain

threats, and to irreversible or persistent impacts over transient impacts.³

- 6.4 However, neither criteria really take into account deep uncertainty on the short- and long-term possible and likely consequences, for better or worse according to disputable values. Even less so do they consider such choices as fuzzy gambles on critical and sometimes fateful stakes. And both criteria and similar ones completely ignore actual choice criteria of high-level decision-makers, such as public support or opposition, political and material profits and losses, short-termism and so on – all further aggravated by multiple biases, including depth psychological ones in addition to simpler ones explored by experimental psychology.
- 6.5 Also usually ignored are cultural impacts; value diversity sensitivity maps, decision delay options combined with structured learning, the legitimate role of pattern-recognizing intuition, and more. Even more amazing is lack of giving due weight to the historic fact that nearly all technologies are not only error-prone, but are also earlier or later used for the worse, such as tribal slaughters and damaging greed.
- 6.6 In short, decision frames fitting real-life high-stake fuzzy gambles within their internal and external contexts are sorely underdeveloped. And second- and third-best criteria, which are available as illustrated above and are much better than nothing, are hardly applied as required – because of short-sighted vested interests and policy inertia. This is clearly demonstrated in the relatively clear-cut case of global climate change.
- 6.7 Ethics theory distinguishes between rule-based ethics, utilitarian ethics, and virtue ethics. In addition to upgraded decision criteria, I propose a similar approach to Singularity choices, including requiring from high-level decision-makers “fuzzy gambling decision virtues,” such as deep uncertainty sophistication, global grand-strategic perspectives, and long-term pondering horizons – within future human evolution frames when Singularity issues are at stake. This leads to the crucial issue who should be the decision makers on major Singularity choices.

7. Global Future-Shaping Super-Elite

- 7.1 It is essential to recognize, however politically incorrect and in many respects disturbing, that unavoidably (until humanity perhaps becomes super-human) a miniscule part of humanity, however constrained, dominates nearly all important future-impacting decisions. To put it into guesstimated orders of magnitude, no more than, say, a maximum of ten

³ Max More and Natasha Vita-More, eds. *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future* (Malden, MA: Wiley, 2013), Kindle Edition, locations 12207-12223.

thousand humans make or meaningfully influence 90 percent of significant future-impacting choices.

- 7.2 In other words, for every 750 thousand humans there is one significant global future shaper. Even if I am wrong by one order of magnitude and there are 100 thousand humans who belong to the global future-impacting super-elite, which is for sure a large exaggeration, still it is one person per 75 thousand humans.
- 7.3 But in fact, the number of major impactors on the future in relation to the Singularity (and other critical choices) is surely much less than ten thousand, leaving us with a striking though not really surprising conclusion:

Improving the decision virtues of, say, one thousand carefully selected persons can significantly upgrade critical Singularity choices. And doing so is not absolutely impossible.

- 7.4 Not less important are social institutions and processes, such as the free market and social media. However, despite being cybernetic self-regulative, they can be redirected by the global future-impacting super-elite, even though they in turn are partly shaped by the social institutions and processes, and changing them may require radical or even revolutionary measures. Therefore, for our purposes, it is correct to focus mainly on the global future-shaping super-elite.
- 7.5 Let me refine the somewhat rough analysis. The global decision-making super-elite is composed of transformative thinkers, the senior staff of international organizations, powerful economic actors, innovative scientists and technologies, a few civic leaders, some military commanders, select mass media moguls, and single outstanding and senior politicians. But not all parts of this super-elite are of equal importance for coping with Singularity issues. Adopting the distinction by David Priestland between societies dominated by merchants, soldiers, and sages,⁴ with the addition of politicians, decisions related to the Singularity are in most societies dominated by merchants (in a broad sense of that term). In a few societies, military R&D elites dominate Singularity-relevant choices, but they too are subject to economic considerations. Scientists and technologists, who can be viewed as knowledge sages, are critical in laying the foundations of the Singularity. But, alas, they depend for resources on merchants or soldiers.
- 7.6 Nominally, politicians are in overall charge, having ultimate formal decision authority over all Singularity-related choices, subject to more or less constitutional limitations. They are also the only part of the global future-impacting super-elite which is value-wise legitimized to make authoritative future-shaping choices, as discussed in Section 8. Therefore:

⁴ David Priestland, *Merchant, Soldier, Sage: A New History of Power* (London: Penguin, 2012).

The decision vices and virtues of political leaders are of potentially dominant importance in shaping the future of the species.

7.7 In fact, many politicians abandon this task, despite its outstanding importance (or because of it) to other decision-makers as long as no storms erupt. This bodes ill for the future unless rectified, because many other parts of the global future-impacting super-elite and most of the more powerful social processes do not fit the requirements of optimal future-shaping choices. Thus, market processes and merchants are driven by the profit motive, which often degenerates into greed; public opinions suffer from an abundance of tribalism compounded by ignorance on the Singularity; soldiers serve mainly tribal images and interests; scientists-technologists are seeking knowledge and status, and they are dominated by merchants and soldiers who control the needed resources. Worst of all, the vast majority of politicians are infected by power considerations, ignorant about science and technology, short-sighted, and subject to social demands and values that do not serve the future, but which most politicians are unwilling to confront or do not know how to overcome or redirect.

7.8 The diagnosis above is somewhat one-sided. There are individual exceptions, such as globally minded political leaders, socially responsible merchants, and some scientist-statesman/women. But they are constrained by widespread tribalism and profit-seeking, as well as ignorant and capricious publics misdirected by the amusement industry, even in highly developed societies. The dreams of the Enlightenment are further away from reality than ever, despite mass education.

7.9 Typical in some important respects is the belated awakening of President Barack Obama, potentially a nearly optimal Singularity decision-maker, to the world as it is. As clearly put in the memoir of one of his senior advisors:

Ambitious legislative activity was out of the question. Abroad, the forces of tribalism and nationalism were building, like tremors before an earthquake.... After years of ... growing tribalism at home and abroad, he had priced in the shortcomings of the world as it is, picking the issues and moments when he could press for the world that ought to be.⁵

7.10 The next step should be exploration of the required decision virtues of political leaders. But before doing so, the legitimation of politicians to make choices impacting on generations to come requires close examination.

⁵ Ben Rhodes, *The World as It Is: A Memoir of the Obama White House* (New York: Random House, 2017), Kindle Edition, 142-143 and 298.

8. Are Politicians As Such Legitimate Singularity-Shaping Decision-Makers?

8.1 According to democratic values, elected politicians seemingly have the right to make future-shaping choices, subject to some constraints and conditions. This also applies, with adjustments, to senior politicians in non-democratic countries, if they enjoy the support of most of their citizens.

Politicians are the strata in charge of making choices impacting on their societies, subject to constraints such as constitutions and public wishes.

8.2 This is an important and obvious postulate, reaffirming the overriding importance of politics. I do not go as far as Aristotle, who regarded politics as the master science on which all social activities depend. But the future of the Singularity is par excellence a domain for which politicians – not scientists, or markets, or social media – carry by definition the ultimate authority and responsibility. Therefore:

The prevailing abandonment by politicians of fateful future-impacting choices to other actors is worse than dereliction of duty – it is an act of treason against their prime duty.

8.3 But there is a major catch:

By itself being a senior democratically elected political leader (or otherwise being elected or selected) is not an adequate legitimation for making choices shaping the fate of future generations.

This counter-conventional and politically very incorrect statement requires justification. In short, democracy is based on the principle that persons and societies are entitled to determine who is entitled to take decisions impacting on them and their minor children. But future generations cannot vote now. Therefore, being democratically elected by one generation as such does not grant legitimacy to making decisions that are likely to impact significantly on multiple future generations.

8.4 In light of this reasoning so far, the issue can now be reformulated:

On a deeper level, the real question is who is entitled to make future evolution-shaping choices on behalf of humankind as a collective deliberative agency? The only justifiable answer is “politicians who have the moral and cognitive qualities needed for making such decision optimally, together with other merit-based global super-elites having fitting qualities, such as select scientists and technologists.”

8.5 I cannot overestimate the broad and deep significance of this conclusion, on the level of both political philosophy and political institutions and practice. It makes havoc of assuming that a democratically elected global parliament

has the right to take decisions shaping the future of human evolution. Such legitimacy can only stem from merit: having the qualities required for optimally steering the future evolution of humanity as increasingly (though surely not completely) possible.

8.6 Therefore:

The theory and practice of democracy and other political regimes have to be reformulated so as to move towards merit regimes, in particular but not only so in respect to significant future human evolution-shaping choices.

However, I will not further develop this radical conclusion in this paper, moving instead to the required merit taking the form of fitting decision virtues, as especially required from political leaders steering the future evolution of the human species in the context of the Singularity.

9. Required Decision Virtues

9.1 The compelling conclusion is demanding:

Senior global future-impactors, and in particular political leaders, require personal decision virtues and supportive environments very different from the prevailing ones.

9.2 As noted, at this stage it is necessary to specify the qualities needed by future-shaping political leaders (and, with adjustments, by other members of the future-shaping super-elite), to enable them to approximate optimal fuzzy gambles on Singularity issues. But in view of the limits of this paper and the two books I have written on that subject,⁶ I limit myself here to a list of twelve representative qualities, in no particular order:

1. commitment to the long-term future of the human species as a priority task, together with efforts to facilitate the thriving of contemporary humans;
2. pondering in terms of the evolutionary processes shaping the human future, as transformed by the Singularity;
3. globalism overriding tribalism, combined with political skills, making doing so feasible;
4. good science and technology literacy, with emphasis on Singularity-related domains;
5. multicultural insights;
6. intense innovation-friendliness;
7. a strong “inner citadel,” combined with the Kantian rule *saper aude* (dare to rely on your own potential abilities);

⁶ *For Rulers*, as mentioned; and *Avant-Garde Politician: Leaders for a New Epoch* (Washington, DC: Westphalia Press, 2014).

8. uncertainty-sophistication;
 9. pronounced reasoning abilities combined with open-ended intuition on Singularity-related issues;
 10. seeking advice from Singularity and human evolution professionals, including especially Singularity policy scientists (as discussed in a companion paper of mine, following), subject to careful screening;
 11. crisis coping skills, with emphasis on utilizing them to implement Singularity-coping ideas that cannot be realized without conservatism-disrupting and mentality-shocking events;
 12. Constant learning and critical self-reflectivity related to major Singularity challenges.
- 9.3 No human being, even with soft enhancement, can be outstanding in all these decision virtues. And, in the foreseeable future reliance on super-intelligent robots to take care of humanity will not become practical; and if and when it becomes available, it will usurp human authority and nullify human moral responsibility for existential choices, while also endangering the future of humanity. But there are enough examples of political leaders, however scarce, who have clearly demonstrated the potential to become adequately qualified to make good future-shaping decisions, without being outstanding in all respects and always arriving at optimal choices.
- 9.4 Furthermore, a well-designed global leadership seminary can help carefully selected participants to develop adequate decision virtues. This brings us to the institutional requirements of composing and implementing well-crafted future-impacting choices, which are far beyond the capacities of contemporary global regimes – but not in the realm of the impossible, given the likelihood of crises breaking the tyranny of the status quo.

10. Enforcement Regimes

- 10.1 Preventing dangerous misuses of Singularity knowledge and tools, whether on purpose or accidentally, is impossible without a radically novel global regime accompanied by painful value transformations, which will be strenuously resisted. Thus, an adequate enforcement regime must be global in scope, overriding state sovereignty, breaking through tribalism, and having forceful instruments to impose its authority when necessary.
- 10.2 Furthermore, it may have to impose personal duties in addition to human rights, regulate and sometimes limit research freedom and technology marketing, engage selectively in intrusive intelligence collection, and be entitled to impose, after due process, harsh punishments. Some property rights, ownership of mass killing weapons, and promotion of hate ideologies will also have to be inhibited. And conflicts that may escalate to catastrophic levels will have to be resolved, with imposed measures and dictated “agreements” as may be necessary.

- 10.3 Taking care of human welfare and fairness will remain the task of national governments and the United Nations with its agencies, subject to quite some reforms. But containing dangers posed by the Singularity on lines illustrated above requires some kind of “Platonic Global Leviathan,” subject to controls and maximum reliance on the subsidiarity principle – but with a preponderance of global enforcement tools.
- 10.4 The concept of global leviathan is clear enough for the limited purposes of this paper for all who are familiar with relevant writings by Thomas Hobbes, though the proposed authority will be less autocratic and more circumscribed than the absolute ruler proposed by Hobbes. But the term “Platonic” needs some explanation.
- 10.5 Taking into account that the quality of an organization depends primarily on the quality of its senior leaders and staff, it is essential that outstanding politicians supported by excellent professionals constitute its human dimension. This returns me to Plato’s *Republic*, which proposed rule by philosophers. Leaving aside the lifestyles dictated by Plato to the rulers, which are both impossible to realize and not fully necessary, the concept of philosopher that was probably in the mind of Plato (though never explicated in his surviving writings) was one of constant seekers of truth in a comprehensive meaning of that term. If we add the ideas of Michel Foucault on truth and power, we have a good basis for conceptualizing the requirements of the heads of the Platonic Global Leviathan, which add to and go beyond and above the qualities required for being a decision-virtuous future impactor, as already discussed. Thus, deep understanding of human evolution, a fusion between idealism and realism, total exclusion of personal considerations in making important fuzzy-gambling decisions, complete disconnection from tribal identity, psychoanalytic measures to reduce mind-distorting depth processes, proven outstanding pattern-recognizing intuition, and some features of a warrior combined with compassion – these illustrate the extra qualities required for heading and running the proposed global authority.
- 10.6 Complementarity of the leaders of the Platonic Global Leviathan can help to achieve emergent synergetic qualities meeting more or less such demanding requirements, which surpass individual potentials. Carefully dosed mind enhancement may also be of much help.
- 10.7 To further illustrate the counter-conventional features of the senior global authority leadership, it may well be advisable to fill many positions by coadoption, so as to strengthen independence. But it is too early to go into such details, which need consideration by outstanding teams rather than by me thinking alone.
- 10.8 I think enough has been said to provide readers with a sense of what is absolutely needed, but also completely impossible given the world as it is.

There is only one way out of this aporia if we want to be realistic as required:

The only way to establish an adequate global authority approximating the features of a Platonic Global Leviathan, which are essential for containing the risks of the Singularity, is to have good designs ready and to prepare needed knowledge and appropriate political leadership and professionals to utilize global major crises for realizing in stages what is essential but impossible without painful creative destruction.

11. Global Crises As Indispensable Opportunities

11.1 To put it bluntly, without the whip of catastrophes, it is extremely unlikely that humanity will do what is essential to assure survival and thriving while undergoing the nearly apocalyptic Singularity metamorphosis. However, harsh transition crises are assured. Thus, explosive social conflicts resulting from high costs of widely desired human enhancement, kitchen-mutated viruses used for mass killings, escaped nanorobots creating global havoc, large-scale unemployment caused by growing use of broadly intelligent robots instead of human workers, mass migration driven by climate change becoming violent when refugees are not admitted into rich countries, persistent global economic crises impoverishing rich and poor alike resulting from molecular engineering – these are just a few of the possible and in part quite likely catastrophe scenarios, with near certainty that at least some of them will occur in the foreseeable future. Therefore:

Paradoxically, serious but not human-survival-endangering global crises provide the best chance for realizing measures without which the long-term thriving and also the existence of the human species is in doubt, given the risks and chances of the Singularity.

11.2 However, enjoying ourselves in what Georg Lukács aptly termed “Grand Hotel Abyss” in the face of species-endangering catastrophes may well be fatal.

Major global crises are an essential condition of realizing what is necessary for preventing fatal events, but they are not sufficient. Crises can just as well result in danger-escalating panic reactions, which assure further and harsher crises up to endangering the survival of the human species.

11.3 Therefore, urgent action is needed to increase the likelihood of utilizing major crises for establishing the needed global regime staffed by outstanding decision-virtuous political leaders supported by extremely qualified Singularity-policy professionals and broad global consensus. This leads us to the concluding pressing question “What is to be done now”?

12. What Is to be Done Now

12.1 Main suggestions for action now, to utilize coming crises for the better and perhaps also to reduce their costs, include for instance, with overlaps:

- Setting up a global think tank network, based on existing centers studying catastrophic dangers as well as select individuals, to work part time and full time in multidisciplinary teams on major Singularity decision issues and composing humanity-craft (a term I derive from statecraft applied to the human species) options.
- Establishing, as mentioned, a global public leadership seminary, dedicated to developing political leaders and professionals with the necessary decision virtues.
- Activating a global scientific council headed by select Nobel laureates and limited to, say, 150 scientists, philosophers, free-floating intellectuals etc., to serve as a kind of scientific senate, and when necessary, science court, as several times proposed, discussing major Singularity dangers, elaborating and applying codes of ethics for Singularity scientists and technologies, and more.
- Building an intranet for open and closed discussion of main Singularity choices by carefully screened participants reflecting different backgrounds, with canvassing of ideas from interested publics at large.
- Bringing together a small group of highly qualified persons, including also former senior political leaders, to work out alternative designs of the needed global authority or adequate alternatives.
- Strengthening informal colleges of concerned social activists and a variety of leaders together with scientists, philosophers, technologists, etc. – to mobilize broad public support for needed measures including an adequate global regime.

12.2 There is quite some activity in such directions, but much more is needed and can and should be done urgently, beyond my limited creativity and the constrained scope of this paper. But one dimension of urgent action which is in multiple ways an important facet of preparing for the Singularity and, even more so, an integral central part of the Singularity, is large-scale and long-distance space exploration. Dr. Bob Krone, who is highly qualified and fruitfully active in this ultra-challenging domain, discusses this possibilities and requirements in the following essay.

Copyright © Yehezkel Dror, 2018. All rights reserved.