An overview of models of technological singularity

Anders Sandberg

Future of Humanity Institute, Oxford University
Suite 8, Littlegate House, 16/17 St Ebbe's Street
Oxford, OX1 1PT
Tel: +44 (0)1865 286279
Fax: +44 (0)1865 286983
anders.sandberg@philosophy.ox.ac.uk

Abstract

This paper reviews different definitions and models of technological singularity. The models range from conceptual sketches to detailed endogenous growth models, as well as attempts to fit empirical data to quantitative models. Such models are useful for examining the dynamics of the world-system and possible types of future crisis points where fundamental transitions are likely to occur. Current models suggest that, generically, even small increasing returns tends to produce radical growth. If mental capital becomes copyable (such as would be the case for AI or brain emulation) extremely rapid growth would also become likely.

Introduction

The set of concepts today commonly referred to as "technological singularity" has a long history in the computer science community, with early examples such as:

> One conversation centered on the ever accelerating progress of technology and changes in the mode of human life, which gives the appearance of approaching some essential singularity in the history of the race beyond which human affairs, as we know them, could not continue. *Ulam, S., Tribute to John von Neumann,* Bulletin of the American Mathematical Society*, vol 64, nr 3, part 2, May, 1958, p1-49.*

and

> 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 ultraintelligent 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. *I.J. Good, (Goo65)*

The unifying theme of these two examples is accelerated technological change leading to a rapid transition to a state where the current human condition would be challenged.

Technological singularity is of increasing interest among futurists both as a predicted possibility in the mid-term future and as subject for methodological debate. The concept is used in a variety of contexts, and has acquired an unfortunately large number of meanings. Some versions stress the role of artifical intelligence, others refer to more general technological change. These multiple meanings can overlap, and many writers use combinations of meanings: even Vernor Vinge's seminal essay (Vin93) that coined the term uses several meanings. Some of these meanings may imply each other but often there is a conflation of different elements that likely (but not necessarily) occur in parallel. This causes confusion and misunderstanding to the extent that some critics argue that the term should be avoided altogether (Tyl). At the very least the term "singularity" has led to many unfortunate assumptions that technological singularity involves some form of mathematical singularity and can hence be ignored as unphysical.

This paper is attempting a simple taxonomy of models of technological singularity, hopefully helping to disambiguate the different meanings of the word. It also aims at a brief review of formal quantitative models of singularity-like phenomena, in the hope of promoting a more stringent discussion of these possibilities.

Definitions of technological singularity

A brief list of meanings of the term "technological singularity" found in the literature and some of their proponents:

A. Accelerating change Exponential or superexponential technological growth (with linked economical growth and social change) (Ray Kurzweil (Kur05), John Smart (Smang))

B. Self improving technology Better technology allows faster development of new and better technology. (Flake (Fla06))

C. Intelligence explosion Smarter systems can improve themselves, producing even more intelligence in a strong feedback loop. (I.J. Good (Goo65), Eliezer Yudkowsky)

D. Emergence of superintelligence (Singularity Institute)[1]

E. Prediction horizon Rapid change or the emergence of superhuman intelligence makes the future impossible to predict from our current limited knowledge and experience. (Vinge, (Vin93))

F. Phase transition The singularity represents a shift to new forms of organisation. This could be a fundamental difference in kind such as humanity being succeeded by posthuman or artificial intelligences, a punctuated equilibrium transition or the emergence of a new metasystem level. (Teilhard de Chardin, Valentin Turchin (Tur77), Heylighen (Hey07))

G. Complexity disaster Increasing complexity and interconnectedness causes increasing payoffs, but increases instability. Eventually this produces a crisis, beyond which point the dynamics must be different. (Sornette (JS01), West (BLH[+]07))

H. Inflexion point Large-scale growth of technology or economy follows a logistic growth curve. The singularity represents the inflexion point where change shifts from acceleration to deacceleration. (Extropian FAQ, T. Modis (Mod02))

I. Infinite progress The rate of progress in some domain goes to infinity in finite time. (Few, if any, hold this to be plausible[2])

The three major groupings appear to be accelerating change, prediction horizon and intelligence explosion leading to superintelligence (as originally noted by (Bos98) and discussed in (Yud07)).

In addition to the general meaning(s), the singularity might be local or global (capability takeoff of an entity or small group, or broad evolution of the whole economy), fast or slow (occuring on computer timescales, hardware development timescales, human timescales, or historical timescales). There is also confusion over whether the salient issue is the point/eventlike charac-

ter, the historical uniqueness, the nature of the overall process or the big historical trend.

For the purpose of this paper I will focus on the growth aspect: accelerating change, self-improving technology, intelligence explosions and the complexity disaster (and to some extent the inflexion point) all involve the growth of technological or cognitive capability. The nature and speed of this growth is important both for understanding what claims are actually made and for considering the implications of such a process, if it takes place.

As noted in (Hey97) models for the technological singularity, whether qualitative or quantitative, should not be taken too literally. Models include what is considered to be the important and relevant features of a system while abstracting away from minor, obscuring features. They are useful because they force us to show our assumptions, bringing them into the open. Their predictions may be less important than demonstrating how changes in assumptions affect outcomes.

Models

Linear takeover (Type D, F)

Singular events can occur when one form of growth outpaces another. This form of "linear singularity" does not necessarily involve any acceleration of progress.

Eliezer Yudkowsky has presented a simple sketch of why apparent progress of AI might be deceptive: the rate of progress is actually high, but starts from a very low level. This means that the rapid development is not easily detectable, until it suddenly passes the relatively narrow human range and goes beyond[3]

A simple formalisation of the model would be that human intelligence X is constant while machine intelligence $Y(t) = a + bt$ grows linearly with time. At present $X \gg Y(t)$. We are less sensitive to differences in less-than-human intelligence than to differences in human-level intelligence; our subjective experience of intelligence (and possibly its ability to affect the world) is a rapidly concave function, for example an exponential curve. The *apparent* intelligence of machines would be growing as e^{a+bt}, with a "surprising" eventual jump from an apparently unchanging low state to rapid growth near human level. This is essentially the same claim made by Kurzweil in regards to how most people do not understand exponential growth, although the

---

[1] "The Singularity is the technological creation of smarter-than-human intelligence", `http://singinst.org/overview/whatisthesingularity`

[2] The exception may be the Omega Point theory of Frank J. Tipler, which predicts an infinite information state reached at the future timelike singularity of a collapsing universe. However, while this (physical) singularity is reached in finite proper time, it is not reached in finite subjective time and there is literally no "after" (BT86).

---

[3] "The best answer to the question, 'Will computers ever be as smart as humans?' is probably 'Yes, but only briefly" *Vernor Vinge*. Samuel Butler made a somewhat similar point in his article "Darwin among the Machines" (*The Press*, Christchurch, 13 June 1863) and novel *Erewhon* (1872).

growth here may not be nonlinear (only its detectability is).

Logistic growth (type H)

It is commonly argued that exponential growth is unsustainable, since it requires resources that will eventually run out. This is true even if space colonization is available, since the volume of space that can be reached in time t grows at most as $4\pi c^3 t^3/3 \propto t^3$, where c is lightspeed. This means that eventually the growth rate will have to decline to at most a polynomial growth rate.

However, if the growth is in the domain of knowledge, culture or pure value it is less obvious that there exists an upper limit or that the growth is strongly resource constrained. There are some physical limits on how much information can be stored within a region of the universe but these are very wide; at most they preclude infinite information storage[4].

There could exist limiting factors to technology based on the (possible) finite complexity of the universe: there would only exist a finite amount of useful facts about nature that can be discovered. As more and more become known the difficulty of finding new ones increases. This would be compatible with a logistic growth equation:

$$X'(t) = r(K - X(t))X(t)$$

where $X(t)$ is knowledge, r is a growth rate and K is the upper threshold of what is knowable. Initially $X(t)$ grows exponentially, but eventually saturates and slows down, asymptotically approaching $X(t) \to K$. Other forms of value might still grow after saturation, but the knowledge will remain bounded.

In this case the singularity would likely denote the inflexion point, the (historically very brief) transition from a pre-technological state to a maximally advanced state.

Metasystem transition (type F)

A metasystem transition is the evolutionary emergence of a higher level of organisation or control in a system (Tur77). A number of systems become integrated into a higher-order system, producing a multi-level hierarchy of control. Within biology such evolutionary transitions have occured through the evolution of self-replication, multicellularity, sexual reproduction, societies etc. where smaller subsystems merge without losing differentiation yet often become dependent on the

---

[4]The Bekenstein bound on the entropy to energy ratio of a system with radius R and total energy E constrains the information stored inside to be less than $2\pi RE/\hbar c \ln(2)$. For a 1 kg, 1 meter system the bound is on the order of 10^{43} bits (Bek81).

larger entity (SS95). At the beginning of the process the control mechanism is rudimentary, mainly coordinating the subsystems. As the whole system develops further the subsystems specialize and the control systems become more effective. While metasystem transitions in biology are seen as caused by biological evolution, other systems might exhibit other forms of evolution (e.g. social change or deliberate organisation) to cause metasystem transitions. Extrapolated to humans, future transitions might involve parts or the whole of the human species becoming a super-organism (TJ89).

As a model for technological singularity the metasystem transition is largely qualitative rather than quantitative. Without a detailed model of the subsystem dynamics the transition is undefined. However, the biological and social examples given by theorists, would give some analogical data on how future transitions could work (e.g. as punctated equilibrium models). A key issue worth exploring may be to what degree such transitions are convergent or divergent, e.g. whether under similar circumstances the transitions would produce similar or different forms of organisation[5].

Accelerated metasystem transition (type A,B,F)

Heylighen argues (Hey07) that the evolutionary forces acting on technology (and other systems) tend to lead to ephemeralization, doing more with less, due to selective pressures in any resource-constrained environment. This produces increasing efficiency in the use, processing and transport of matter, energy and information, reducing "friction" and enabling longer controllable causal chains. Similarly innovations in institutions, mediators and stigmergy (indirect influences on other agents through this environment, "the medium") produces increasingly global coordination. He argues that stigmergy accelerates evolution, since already found solutions spread to other agents: improvement of the medium facilitates further innovation, which helps improve and spread the medium. He compares this to the population growth model of Korotayev (Kor07) where the population N is controlled by a logistic growth equation where the carrying capacity is proportional to the overall productivity T of technology:

$$N'(f) = a(bT - N)N$$

Technology grows proportional to population and technology:

$$T'(t) = cNT$$

This produces a hyperbolic growth curve with a finite time singularity. Heylighen notes that the deviation in population growth from hyperbolic growth in the 1970's is still compatible with the general model if we interpret it as a shift from an r-strategy (fast reproduction, short life) to a K-strategy (slow reproduction, long life). In

---

[5]This is one of the concerns of the theory in (Smang)

a K-strategy more human capital is invested in the off-spring, and they will provide more back, making the T factor of the second equation more important than the N factor. He predicts that the above trends will lead to the formation of a global superorganism of some kind over a relatively short timespan.

Accelerating change

Economic input output models (type A)

Input-output models depict the economy as A, a matrix denoting the number of units a sector of the economy needs to buy from another sector to produce a unit of its own output, and B, a matrix of how many units from other industries are needed for the production itself (e.g equipment, buildings etc). If the output levels of different sectors at time t are $X(t)$ the amount of goods delivered to households and other final users will be (Leo86):

$$Y(t) = X(t) - AX(t) - B[X(t+1) - X(t)]$$

If all the outputs are re-invested into the economy $Y(t) = 0$ and the equation produces a dynamical system

$$X(t+1) = [B^{-1}(1 - A - B)]X(t)$$

The growth will be exponential, with a rate set by the largest eigenvalue of the bracketed matrix and a production vector X tending towards the eigenvector corresponding to the value. If the consumption $Y(t)$ is nonzero the generic case is still exponential growth: matrix recurrences of the form $X(t+1) = CX(t) + d$ or differential equations like $X'(t) = CX(t) + d$ have solutions tending towards $\Lambda e^{\lambda t}$ if C is diagonalizable.

Endogenous growth models (type A,B,I)

Endogenous growth theory models the growth of an economy with improving technology, where the technology is assumed to be growing as a function of the economy and allocation of resources to it. It was developed as a response to exogenous growth models, where diminishing returns predict that growth would stop rather than continue.

> ...while the part which nature plays in production shows a tendency to diminishing return, the part which man plays shows a tendency to increasing return. The law of increasing return may be worded thus:An increase of labour and capital leads generally to improved organization, which increases the efficiency of the work of labour and capital. *Marshall (1920), Book 4, Chapter XIII, Paragraph IV.XIII.11.*

However, published models usually appear to try to avoid "unrealistic" explosive growth, by selecting equations for the growth of knowledge that preclude it. For example, (HI04) amend their model not only to avoid negative knowledge but also to preclude superexponential growth. In a later paper (HI07) they state:

> We take the view that a plausible description of the evolution of knowledge should satisfy two asymptotic conditions. Looking *forwards*, we follow Solow (2000) in maintaining that infinite knowledge in finite time is impossible. Looking *backwards*, we require knowledge to vanish in the infinite past, but not in finite time. We call an evolution *plausible* if it satisfies these criteria.

This paper shows that for knowledge growth

$$X'(t) = gX(t)^\phi \qquad (1)$$

knowledge is either zero at all times, or increasing over time. In the latter case if $\phi < 1$ the knowledge starts from zero in the finite past, if $\phi > 1$ knowledge diverges to infinity in the finite future, or if $\phi = 1$ knowledge is always finite and grows exponentially without any singularities. It hence concludes that only the exponential case is possible. Adding (endogenous or exogenous) population growth does not help in the generic case.

However, a simple model predicting a finite-time singularity might still be relevant if other, unmodelled, factors near the potential singularity prevents it from actually occuring. The assumption that a model must hold for an infinite timespan (especially when it is a highly abstracted model of a complex process) seems unwarranted. Models exhibiting finite-time singularities can be a quantitatively or qualitatively accurate fit to current or near-future growth, although they eventually leaves the domain of applicability.

As noted in (JS01):

> Singularities are always mathematical idealisations of natural phenomena: they are not present in reality but foreshadow an important transition or change of regime. In the present context, they must be interpreted as a kind of "critical point" signaling a fundamental and abrupt change of regime similar to what occurs in phase transitions.

The observation that finite-time singularities are generic when economies of scale in knowledge production exist ($\phi > 1$) seems to support further study of superexponential growth, and cast doubt on the assumption that knowledge growth should always be modeled as exponential.

A widely quoted growth model exhibiting finite-time singularity is the model of Kremer (M93): the total population is $L(t)$ and total economic output $Y(t)$, and their ratio $Y(t)/L(t) = \bar{y}$ is set to a subsistence level $\bar{y}$ which is fixed. Output depends on technology/knowledge $A(t)$ and labour proportional to $L(t)$:

$$Y(t) = Y_0[A(t)L(t)]^{1-\alpha}$$

where $0 < \alpha < 1$. The growth rate of technology is assumed proportional to population and technology level:

$$A'(t) = BL(t)A(t)$$

with the explicit assumptions that larger populations have more talented people that can advance technology and that new technology largely is obtained by leveraging existing technology (hence, it implies a type B singularity). Combining the two equations produces

$$L'(t) = [B(1 - \alpha)/\alpha]L(t)^2$$

Hence the population (and technology) has a finite time singularity.

(SS96) further discuss multivariate extensions of the model with capital $K(t)$ producing output

$$Y(t) = [(1 - a_K)K(t)]^\alpha [A(t)(1 - a_L)L(t)]^{1-\alpha}$$

where a_K is the fraction of capital stock used in R&D, a_L the fraction of labour used in R&D. Innovation is produced as

$$A'(t) = B[a_K K(t)]^\beta [a_L L(t)]^\gamma [A(t)]^\theta$$

for positive constants B, β and γ. Assuming exogenous saving rate s and no depreciation gives

$$K'(t) = sY(t) = s[(1 - a_K)K(t)]^\alpha [A(t)(1 - a_L)L(t)]^{1-\alpha}$$

This system can achieve finite time singularities even for fixed population if $\theta + \beta > 1$. Past innovation and capital can create explosive growth even when each of the factors in isolation cannot.

Population-technology model (Type A,F,I)

A population theoretic model similar to endogenous growth was formulated by Taagepera (Taa79). It links the population P with technology T and nonrenewable resources R via a modified logistic growth model:

$$P'(t) = k_0 \left[1 - e^{-aT^n}\right]\left[1 - \frac{(R + C)P}{fT^q R}\right]P$$

where the first factor represents growth rate (increased by technology towards a maximal rate k_0). Technology increases as a power of population size for small populations and independent of it at large sizes:

$$T'(t) = h\left[\frac{P}{U + P}\right]^m T$$

where U is a characteristic population size. The rate of depletion of resources was either modeled with a complete recycling model, where advanced societies can recycle most material

$$R'(t) = -\frac{fVTP}{(V + T)^2}$$

where V is the critical technology level where resource depletion levels out. The other possibility was merely a stabilization of per capita depletion rate

$$R'(t) = -\frac{fTP}{V + T}$$

For the initial state $P \ll U$, $T \ll V$, $T^n \ll 1/a$ and $R \gg C$ and the equations produce hyperbolic growth. Depending on whether the first, the middle two or the last inequalities change sign first there are three outcomes. The first case, resource depletion, produces an eventual saturation population with a likely eventual decrease. In the middle case recycling or per capita stabilisation becomes significant before population is large and resources has run out, and hyperbolic growth continues until one of the other cases occurs. If the population becomes large but there are still resources population growth is doubly exponential.

Taagepera notes that the model while avoiding finite time singularities still has crises where it shifts from one mode to another, possibly over very short periods of time.

Law of Accelerating returns (type A,B)

Ray Kurzweil formulates the "law of accelerating returns" as (Kur01):

- Evolution applies positive feedback in that the more capable methods resulting from one stage of evolutionary progress are used to create the next stage. As a result, the rate of progress of an evolutionary process increases exponentially over time. Over time, the "order" of the information embedded in the evolutionary process (i.e., the measure of how well the information fits a purpose, which in evolution is survival) increases.

- A correlate of the above observation is that the "returns" of an evolutionary process (e.g., the speed, cost-effectiveness, or overall "power" of a process) increase exponentially over time.

- In another positive feedback loop, as a particular evolutionary process (e.g., computation) becomes more effective (e.g., cost effective), greater resources are deployed toward the further progress of that process. This results in a second level of exponential growth (i.e., the rate of exponential growth itself grows exponentially).

- Biological evolution is one such evolutionary process.

- Technological evolution is another such evolutionary process. Indeed, the emergence of the first technology creating species resulted in the new evolutionary process of technology. Therefore, technological evolution is an outgrowth of–and a continuation of–biological evolution.

- A specific paradigm (a method or approach to solving a problem, e.g., shrinking transistors on an integrated circuit as an approach to making more powerful computers) provides exponential growth until the method exhausts its potential. When this happens, a paradigm shift (i.e., a fundamental change in the approach) occurs, which enables exponential growth to continue.

Kurzweil models the growth as composed of a velocity of technology growth $V(t) = c_1 W(t)$ driven by and driving the total world knowledge $W'(t) = c_2 V(t)$ (where c_1 and c_2 are constants). This has an exponential solution with growth rate $c_1 c_2$. He then assumes that there are exponentially increasing resources for computation $N(t) = c_3^{c_4 t}$ and that world knowledge grows proportional to the product of $V(t)$ and $N(t)$: $W'(t) = c_1 V(t) N(t) = c_1 c_2 c_3^{c_4 t} W(t)$ which has a double exponential solution of the form $W(t) \propto \exp(Ct^2)$.

It is not clear why the exogenous computational resources $N(t)$ are not increasing proportional to $W(t)$, for example if the economy and population providing them were positively influenced by knowledge growth, rather than just growing by a steady exponential. If there were any positive feedback with $W(t)$ the equations would show essentially the same behavior as equation 1 and have a finite time singularity.

Vinge/Moravec model (type A,B,I)

The original singularity essay by Vinge (Vin93) does not describe any quantitative model. It says:

> "When greater-than-human intelligence drives progress, that progress will be much more rapid. In fact, there seems no reason why progress itself would not involve the creation of still more intelligent entities – on a still-shorter time scale. The best analogy that I see is with the evolutionary past: Animals can adapt to problems and make inventions, but often no faster than natural selection can do its work – the world acts as its own simulator in the case of natural selection. We humans have the ability to internalize the world and conduct "what if's" in our heads; we can solve many problems thousands of times faster than natural selection. Now, by creating the means to execute those simulations at much higher speeds, we are entering a regime as radically different from our human past as we humans are from the lower animals."

This is a re-statement of the intelligence explosion model of Good (Goo65). Vinge also argues that current technological progress (especially in computing) is influenced by automation of design, producing strong incentives to further automate it. This leads to a strong feedback where either AI, intelligence amplification or distributed intelligence produces an intelligence explosion. A formalisation of this would be to model technological improvement as an exponential process, where new technology reduces the doubling time. However, whether this produces a finite time singularity or merely accelerating growth depends on the exact functional form.

Hans Moravec explored a model based on Vinge (Mora; Morb). He initially assumed that "world knowledge" $X(t)$ produces an exponential speedup of computer performance $V(t) = e^{X(t)}$. In the case of a constant number of humans working unassisted the growth of knowledge would be linear, $X'(t) = 1$, producing $X(t) = t$ and $V(t) = e^t$ (standard Moore's law).

If knowledge growth is instead driven directly by computers $X'(t) = V(t) = e^{X(t)}$, producing the solution $X(t) = \log(-1/t)$ ($t < 0$). This corresponds to a slow growth reaching a singularity. A further model added computer power to human power, giving $X(t) = \log(1/(e^{-t} - 1))$ which has linear growth for $t \gg 0$, becomes roughly exponential near $t = -1$ and has a singularity at $t = 0$.

Observing that $V(t) = e^{X(t)}$ was likely far too optimistic he then went on to show a variety of models where $V(t)$ was a concave increasing function of $X(t)$ produce finite time singularities. These results are substantially the same as the observation about equation 1. Slightly superlinear growth, e.g. $X'(t) = (1 + \log(X(t)))X(t)$ merely produces fast double exponential growth, $X(t) = \exp(e^t - 1)$, but squared logarithmic growth $(X'(t) = (1 + \log^2(X(t)))X(t))$ gives rise to a true singularity $(X(t) = \exp(\tan(t)))$. In fact, if the logarithm has exponent > 1 there is an eventual mathematical singularity.

In a subsequent discussion between Moravec, Kurzweil, and Vinge (KVM) Kurzweil noted:

> My sense is that it is difficult to justify on theoretical grounds having the rate of increase of knowledge be equal to the "size" of knowledge raised to a power greater than 1.
>
> However, I do think it is feasible to justify separate components of growth, some (or one) of which have an exponent of one, and some (or one) of which have a log exponent. The analysis above points to a log exponent. I believe we can justify that the "value" of a typical network or of a database does not expand proportional to its size, but to the log of its size. For example, the value of the Internet to me does not double if the number of users double. Rather, its value to me increases in a logarithmic fashion. The exponent of 1 comes from the number of users themselves. Having twice as many users means it is serving twice as many people. So the overall value of the network = $n * \log n$ (n people served times the value of the network to

each person $= \log n$). This varies from Metcalfe's formula that the value of a network $= n^2$.

He suggested that empirical data might be enough to justify a double exponential, but found the unlikelihood of infinite information strong enough to tentatively rule out faster increments.

Solomonoff (type A,B,I)

Ray J. Solomonoff presented a model of progress in AI based on the idea that once a machine is built that has general problem solving capacity near that of a human in an area such as industry or science such machines will be used for in those areas, speeding up technological progress. This would eventually lead to construction of machines with capacities near or beyond the original computer science community (Sol85). In particular, he notes that the key factors is that machine intelligence will eventually become cheaper than human intelligence, and that high initial training costs can be offset by copying already trained machines.

His formal model consists of the growth of the effective computer science community (humans plus AI, humans assumed to be constant) $C(t)$ as

$$C'(t) = Rx(t)$$

where R is the rate of money spent on AI and $x(t)$ is the amount of computing power per dollar. $x(t)$ is assumed to grow exponentially with a doubling time inversely proportional to the size of the computer science community:

$$(\log(x(t)))' = AC(t)$$

where A is a constant. Combining equations and assuming $c = x = 1$ at $t = 0$ produces

$$C'(t) = A(C(t)^2 - 1)/2 + R$$

which has a finite time singularity for some t. Assuming a doubling time of x every four years produces $A = \log(2)/4 \approx 0.1733$; with $R = 1$ the singularity occurs in 4.62 years, for $R = 0.1$ in 11.11 years and $R = 0.01$ in 21.51 years.

The model implicitly sets $t = 0$ at the point where the milestone of near-human intelligence is reached; before this computers are not assumed to expand the effective computer science community very much. It is hence a model of just a particular technological community rather than general technological growth, although it is plausible that the rapid increase of $x(t)$ will influence society profoundly.

Hamacher (Type E)

Hamacher commented on the Moravec/Vinge/Kurzweil model, noting that its macroeconomic coarseness ignored issues of coordination problems, competition, re-source allocation and sociological issues. Instead he introduced a cobweb-model of price determination, where supply and demand iteratively determine quantity and price of a product. If $S(q)$ is the supply and $D(q)$ the demand at quantity q, the available price p_t and quantity q_t are updated iteratively

$$p_{t+1} = D(q)$$

$$q_{t+1} = S^{-1}(p_{t+1})$$

Such models have a rich dynamics corresponding to iterated 1-dimensional maps, including stable fixed points, limit cycles and chaotic attractors depending on the choice of $S(q)$ and $D(q)$. In particular the entropy or Lyapunov exponent of the time series q_t produces a prediction horizon: initial uncertainties about the system state grow exponentially at a rate set by these values. As parameters are varied to mimic substitution dynamics the probability of encountering a prediction horizon increase (Ham06).

City economics (Type A,G)

(BLH$^+$07) analysed the economies of scale of cities, showing that many quantities reflecting wealth creation and innovation scale with population with an exponent $\beta > 1$, implying increasing returns. Infrastructure quantities on the other hand scale with $\beta < 1$, showing economies of scale. A larger city will hence produce more per capita but have lower maintenance costs per capita. This is in direct contrast to biological organisms, where the "economy" (heart rate, metabolism, pace of life) decreases with size.

The authors suggested an urban growth equation. Growth is constrained by the availability of resources and their rate of consumption. The resources Y are used both for maintaining the existing infrastructure (at a cost of R per unit time and unit of infrastructure) and to expand it (at a cost of E to get one unit). The allocation can be expressed as $Y = RX(t) + EX'(t)$ where $X'(t)$ is the growth rate of the city. The total growth is

$$X'(t) = (Y_0/E)X(t)^\beta - (R/E)X(t)$$

which has solution ($\beta \neq 1$; for $\beta = 1$ the solution is exponential):

$$X(t) = \left[Y_0/R + (X(0)^{1-\beta} - Y_0/R)e^{-R(1-\beta)t/E} \right]^{1/(1-\beta)}$$

For $\beta < 1$ growth is eventually limited and approaches an asymptotic level, very similar to biological growth. For $\beta > 1$ growth becomes superexponential and eventually reaches a finite time singularity if there are enough resources.

However, the authors do not consider this singularity to be reachable. They argue that eventually there is a transition to a resource-limited state and instead the

growth reverses and eventually collapses. In order to avoid this crisis the system must "innovate", changing the constraining resource limits to allow further growth. By repeating this in multiple cycles collapse can be postponed. The model predicts that the time between cycles decreases as $1/X(0)^{\beta-1}$ where $X(0)$ is the size at the start of a cycle. This produces superexponential average growth as long as sufficiently strong new innovations can be supplied at an ever accelerating rate - until a finite time singularity is reached.

While the model describes cities, the general structure seems applicable to economic systems with increasing returns.

Hanson (Type A)

Robin Hanson has examined the economics of technological singularity using standard economic tools.

He analysed a simple model of investment in the context of technological singularity (Han98b). He found that the curve of supply of capital has two distinct parts, producing two modes of economic growth. In the normal slow growth mode rates of return are limited by human discount rates. In the fast mode investment is limited by the total amount of wealth - the returns are so great that the savings rate become very high. As technology develops the demand for capital slowly increases, nudging the system towards the fast mode (which, given past growth of savings rate, might occur somewhere "near the year 2150"). However, this mode requires fine-tuning the savings rate and the fraction of investment return.

In most senses of technological singularity the amount of available "mental capital" (humans or machines able to do skilled work) increases significantly. Hanson models this in an exogenous growth model that examines economic growth given machine intelligence (Han98a). As machines become more capable they no longer just complement human labor (which tends to increase the demand for skilled labor) but can substitute for it. In the model, human-substituting machines have a falling price, but originally computer investments only buy non-intelligent complementary computers since the price of AI is too high. As the price falls AI starts to replace humans, and human wages will fall along with the computer price.

Introducing machine intelligence makes annual growth rate at least an order of magnitude larger than the current growth rate, with doubling times measured in months rather than years. Partially this is due to the assumption that computer/AI technology grows faster than other technology, partially because machines can be manufactured to meet production demands while humans have to grow up. This also produces a very fast increase in the population of intelligences, which

may rapidly reach Malthusian resource limits: Hanson predicts that per-intelligence consumption will tend to fall. Whether this would lead to increasing per-capita income for humans depends on whether they retain a constant fraction of capital.

Model variations such as endogenous growth, changing work hours, a continuum of job types where some are more suitable for humans than others, and distinguishing human capital, hardware, and software from other capital do not change the essential conclusions.

Hanson found that the transition from a human-dominated to a machine dominated economy could be fast (on the order of a few years, assuming Moore's law-like computer cost development). This model is essentially an economic model of a "linear singularity" (section) where AI passes the human intelligence (or economic efficiency) range. It does not assume technological progress to speed up as the economy grows faster, which, if included, would speed up the transition.

Similar economic effects appear likely to occur if brain emulations (simulations of human brains with sufficient resolution to produce human-equivalent behavior and problem-solving (SB08)) could be created (Han94; Han08a). The presence of copyable human capital leads to a software population explosion limited by the cost and availability of the hardware necessary to run the emulations, producing a high growth rate but diminishing individual wages.

Empirical estimates

Empirical estimates of technological singularity consists of attempts to collate historical (sometimes even paleontological and cosmological) data to estimate whether the "rate of change" is increasing exponentially or superexponentially. This presupposes that the singularity is due to an ongoing, large-scale process in the present or earlier. Intelligence explosion and prediction horizon models can likely not be supported or disputed using this kind of data.

Technological growth (Type A,B,H)

There is a sizeable literature on the exponential growth of scientific resources and output, as well as growth within many technological fields. Similarly there exist various forecasts of imminent ends of progress, either due to resource depletion or more subtle issues.

Rescher's law states that the cumulative output of first-rate findings in science is proportional to the logarithm of how much has been investment, producing merely linear returns as investment grows exponentially (second-rate findings, on the other hand, may grow proportion-

ally to investment and does not necessarily have negligble utility) (WD01).

On the other hand, Meyer and Vallee argued that technological growth in many domains (such as power, manufacturing, speed, typesetting etc.) has been superexponential, likely hyperbolic and composed of smaller logistic growth curves for individual technologies. Since the authors did not assume infinite growth to be feasible, they predicted that the shift from hyperbolic growth to some other regimen would either be through some form of "soft regulation" or a catastrophe (MV75).

Population (Type A,G,I)

Various authors have fitted human population to a hyperbolic growth curve $N(t) = C/(T - t)$, where population reaches infinity at some time T. This was claimed to be a good empirical fit (vMA60; Taa79; JS01; Kor07), but since 1970 the observed population no longer fits the growth curve. (Kap06) argues that the model is still essentially valid (growth is driven by the square of population) but can be extended and the singularity removed by taking into account finite human reproduction; the changed model instead predicts the year 2000 as the inflexion point of global population growth.

Kremer tested the prediction of his economic growth model, finding that population growth rate historically has been proportional to population. Since the number of technological innovations (which increase the maximum possible population) is assumed to be proportional to population this would give some evidence for hyperexponential growth (M93).

Sequence of economical growth modes (Type A, F, H)

Robin Hanson examined long-term growth of the human economy (Han98c), based on empirical estimations of the past world product. He found that the world product time series over the last two million years could be fit well with a small number of exponential growth modes. One interpretation was as a sum of four exponentials added together[6], but a better fit (with the same number of parameters) was a model with constant elasticity of substitution betwee three exponential growth modes. These modes appear to correspond to the hunting, farming and industrial eras, and had doubling times of 224,000, 909 and 6.3 years. There might also be some evidence for an earlier mode based on brain evolution.

During each time period the economy tends to be dominated by only one mode, and the rate of growth is roughly constant[7]. Between these modes a rapid (faster than the typical doubling time of the previous mode) shift occurs to a mode with a growth rate about two orders of magnitude larger. This is driven by the appearance of new technologies that change the structure and scale of the economy (Han08b).

Given the past appearance of such new modes Hanson proposed that a new mode might occur in the relatively near future. If the number of doubling times during a mode is roughly constant this would place the transition somewhere during the 21st century.

This model involves type A, F and H singularities. Technically, it does not involve any logistic limitation on growth, but rather a logistic-like limitation of increase of growth rate within each mode. It is also unique in predicting multiple *past* singularities (in the sense of type F radical phase transitions).

Sornette (Type A,F,G)

(JS01) fit power laws of the form $(T-t)^\beta$ to world population, GDP and financial data. β is allowed to be complex, implying not only a superexponential growth as time T is approached (due to the real part of the exponent) but also increasingly faster oscillations (due to the imaginary part). The use of this form is motivated with analogy with physics, for example cascades of Rayleigh-Taylor instabilities, black hole formation, phase separation and material failure which all show log-periodic oscillations before the final singularity. Theoretically this has been motivated by considerations of discrete scale invariance and complex critical exponents in non-unitary field theories (SS96).

Their conclusion is:

> The main message of this study is that, whatever the answer and irrespective of one's optimistic or pessimistic view of the world sustainability, these important pieces of data all point to the existence of an end to the present era, which will be irreversible and cannot be overcome by any novel innovation of the preceding kind, e.g., a new technology that makes the final conquest of the Oceans and the vast mineral resources there possible. This, since any new innovation is deeply embedded in the very existence of a singularity, in fact it *feeds* it. As a result, a future transition of mankind towards a qualitatively new level is quite possible.

---

[6] A hyperbolic fit produced a worse result.

[7] This contradicts the claim that innovation and growth have been accelerating recently.

Paradigm shifts (type A,F)

Various more or less empirical attempts to predict the rate of paradigm shifts have been made. The level of evidence and modeling is variable. Some examples:

Gerald S. Hawkins proposed the "mindstep equation"

$$M_n = -30,629 + 98,048(1/4 + 1/4^2 + 1/4^3 + \ldots + 1/4^n)$$

for the years where a "mindstep" (a fundamental human paradigm shift) would occur. It converges towards 2053, and is based on 5 (!) data points (Haw02).

Nottale et al. have argued that the main economic crises of western civilization show a log-periodic accelerating law with a finite time singularity ≈ 2080 (NCG01).

Theodore Modis collected a number of evolutionary turning points from a variety of sources, estimating their "complexity jumps" (based on the assumption that the change in complexity would be inversely proportional to the distance to the next) (Mod02). He found that the best fit to complexity growth was a large-scale logistic curve, placing the present about halfway through the history of the universe. Ray Kurzweil re-plotted the same data and concluded that the paradigm shift rate is doubling every decade, producing an extremely rapid rate of acceleration (Kur01). This led to a spirited response critiquing the independence of the data, Kurzweil's methodology and the general assumption of accelerating technological growth (Mod06).

D.J. LePoire attempted to estimate the progress of physics by assigning events to different subfields (e.g. optics, wave theory, quantum mechanics etc.) and plotting curves of the fraction of results in the subfield that has been completed at a certain time; these were roughly logistic. Using their median times, the assumption that overall physics progress overall follows a logistic curve and that some subfields form natural "stages" produced a fit with center on 1925 (LeP05). His model suggest that the singularity is already past, at least in physics (a strong type H claim).

Rober Aunger has argued that thermodynamics represents a key factor in changing the organization of systems across history, and focused on the emergence of new mechanisms of control of energy flow within systems. Using a dataset of candidates he found an increasing trend of energy flow density and a power law decrease of gap length between transitions. Although he predicted the next transition to start near 2010 and to last 20-25 years, he argued that there has been a plateu in transition lengths for the last century that would preclude a technological singulariy (Aun07)[8].

---

[8]Eric Chaisson makes a similar claim over the past for the rise in free energy rate density (Joules of free energy pro-

Discussion

Generically, mathematical models that exhibit growth tend to exhibit at least exponential growth since this is the signature of linear self-coupling terms. If there exist efficiencies of scale introducing nonlinearities superexponential growth or finite time singularities appear to be generic. Hence it is unsurprising that there is a plethora of models producing rapid or singular growth. The real issue is whether models with a strong mapping to reality can exhibit such phenomena.

It should also be stressed that even if a model admits mathematical singularities its applicability or plausibility may not diminish. Consider Newtonian and Einsteinian gravity, where mathematical singularities occur, yet merely indicate the limits of the applicability of the theory. Mathematical singularities in a growth model are likely indicators for transitions to other domains of growth or that other, unmodeled factors will become relevant close to the point. The plausibility (but obviously not the applicability) of a model is more dependent on its fit to non-singular dynamics in the past and present than its potential misbehavior in the future.

Empirical fits based on "paradigm shifts" or other discrete events suffer from limited data as well as biasing memory effects - we have more available data for recent times, and events close to us may appear more dramatic. It is also easy to cherry-pick data points consciously or unconsciously, fitting expectations (Sch06). In addition there is a risk of biased model selection: people inclined to believe in a singularity of a particular type are more likely to attempt to fit models with that kind of behavior than people resistant to the idea.

Population models may be less biased but do not necessarily track the technological factors of interest directly and are obviously limited by biological limitations. Still, the intersection between population modeling, empirical studies of technological and economic growth and growth models appears to afford much empirically constrained modeling. The fact that long-term (at least) exponential growth can occur and that many not too implausible endogenous growth models can produce radical growth appears to support some forms of the singularity concept.

A common criticism is that technological singularity assumes technological determinism. This appears untrue: several if not all of the singularity concepts in the introduction could apply even if technology just exhibited trends driven by non-deterministic (but not com-

---

cessed per second per kilogram of matter) (Cha98). However, his data on free energy rate does not seem to suggest a curve steadily convergent to a near-future singularity, just a universal trend towards higher rates over time where human technology has increased the slope significantly.

pletely random) microlevel decisions. After all, population models are not criticised as being "biologically deterministic" if they do not model microscale dynamics – ignoring contingency and statistical spread are merely part of the normal modeling process, and further models may if needed include such subtle details. Future models of technological singularity may very well include the multiform microscale or cultural details that shape large-scale progress, but understanding the major structure of the phenomenon needs to start with simple and robust models where the impact of different assumptions is laid bare. It might turn out that technological progress has sufficient degrees of freedom to deviate from any simple model in unpredictable ways, but past trends appear to suggest that large-scale technological trends often are stable.

The most solid finding given the above models and fits is that even small increasing returns in a growth model (be it a model of economics, information or system size) can produce radical growth. Hence identifying feedback loops with increasing returns may be a way of detecting emerging type A singularities.

Endogenous growth models and Robin Hanson's models also strongly support the conclusion that if mental capital (embodied in humans, artificial intelligence or posthumans) becomes relatively cheaply copyable, extremely rapid growth is likely to follow. Hence observing progress towards artificial intelligence, brain emulation or other ways of increasing human capital might provide evidence for or against type A singularities.

There is a notable lack of models of how an intelligence explosion could occur. This might be the most important and hardest problem to crack in the domain of singularity studies. Most important since the emergence of superintelligence has the greatest potential of being fundamentally game-changing for humanity (for good or ill). Hardest, since it appears to require an understanding of the general nature of super-human minds or at least a way to bound their capacities and growth rates.

The dearth of models of future predictability is less surprising: available evidence show that human experts are usually weak at long-term forecasting even without singularities.

Acknowledgments

References

Robert Aunger. A rigorous periodization of 'big' history. *Technological Forecasting and Social Change*, 74:1164–1178, 2007.

Jacob D. Bekenstein. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Phys. Rev. D*, 23(2):287–298, Jan 1981.

Luís M. A. Bettencourt, José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B. West. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences*, 104(17):7301–7306, April 2007.

Nick Bostrom. Singularity and predictability. `http://hanson.gmu.edu/vc.html#bostrom`, 1998.

J.D. Barrow and F.J. Tipler. *The Anthropic Cosmological Principle*. Oxford University Press, 1986.

Eric J. Chaisson. The cosmic environment for the growth of complexity. *Biosystems*, 46(1-2):13 – 19, 1998.

Gary William Flake. How i learned to stop worrying and love the imminent internet singularity. In *Proceedings of the 15th ACM international conference on Information and knowledge management*, page 2, 2006. Arlington, Virginia, USA.

I.J. Good. Speculations concerning the first ultraintelligent machine. *Advances in Computers*, 6, 1965.

Kay Hammacher. Accelerating changes in our epoch and the role of time-horizons. In Vladimir Burdyuzha, editor, *The Future of Life and the Future of our Civilization*, volume III. Springer, 2006.

Robin Hanson. If uploads come first: the crack of a future dawn. *Extropy*, 6(2), 1994. `http://hanson.gmu.edu/uploads.html`.

Robin Hanson. Economic growth given machine intelligence. `http://hanson.gmu.edu/aigrow.pdf`, 1998.

Robin Hanson. Is a singularity just around the corner? what it takes to get explosive economic growth. *Journal of Evolution and Technology*, 2, 1998. `http://hanson.gmu.edu/fastgrow.html`.

Robin Hanson. Long-term growth as a sequence of exponential modes. `http://hanson.gmu.edu/longgrow.pdf`, 1998.

Robin Hanson. Economics of brain emulations. In Peter Healey and Steve Rayner, editors, *Unnatural Selection - The Challenges of Engineering Tomorrow's People,*, pages 150–158. EarthScan, London, 2008.

Robin Hanson. Economics of the singularity. *IEEE Spectrum*, pages 37–42, June 2008.

Gerald S. Hawkins. *Mindsteps to the cosmos*. World Scientific, 2002.

Francis Heylighen. The socio-technological singularity. http://pespmc1.vub.ac.be/SINGULAR.html, 1997.

Francis Heylighen. Accelerating socio-technological evolution: from ephemeralization and stigmergy to the global brain. In George Modelski, Tessaleno Devezas, and William Thompson, editors, *Globalization as an Evolutionary Process: Modeling Global Change*. Routledge, London, 2007.

Hendrik Hakenes and Andreas Irmen. Airy growth - was the take-off inevitable?, 2004.

Hendrik Hakenes and Andreas Irmen. On the long-run evolution of technological knowledge. *Economic Theory*, 30:171–180, 2007.

Anders Johansen and Didier Sornette. Finite-time singularity in the dynamics of the world population, economic and financial indices. *Physica A*, 294:465–502, 2001.

Sergey P. Kapitza. *Global population blow-up and after: the demographic revolution and information society.* Report to the Club of Rome, 2006.

Andrey Korotayev. Compact mathematical models of world system development, and how they can help us to clarify our understanding of globalization processes. In George Modelski, Tessaleno Devezas, and William Thompson, editors, *Globalization as an Evolutionary Process: Modeling Global Change*, pages 133–161. Routledge, London, 2007.

Raymond Kurzweil. The law of accelerating returns. http://www.kurzweilai.net/articles/art0134.html, March 7 2001.

Raymond Kurzweil. *The Singularity Is Near: When Humans Transcend Biology.* Viking Penguin, 2005.

Ray Kurzweil, Vernor Vinge, and Hans Moravec. Singularity math trialogue. http://www.kurzweilai.net/meme/frame.html?main=/articles/art0151.html.

Wassily W. Leontief. *Input-output economics.* Oxford University Press, 2nd edition, 1986.

David J. LePoire. Application of logistic analysis to the history of physics. *Technological Forecasting and Social Change*, 72:471–479, 2005.

Kremer M. Population growth and technological change: One million b.c. to 1990. *Quarterly Journal of Economics*, 108:681–716, 1993.

Theodore Modis. Forecasting the growth of complexity and change. *Technological Forecasting and Social Change*, 69:377–404, 2002.

Theodore Modis. Discussion. *Technological Forecasting and Social Change*, 73(2):104 – 112, 2006.

Hans Moravec. Simple equations for vinge's technological singularity. http://www.frc.ri.cmu.edu/~hpm/project.archive/robot.papers/1999/singularity.html.

Hans Moravec. Simpler equations for vinge's technological singularity. http://www.frc.ri.cmu.edu/users/hpm/project.archive/robot.papers/2003/singularity2.html.

Francois Meyer and Jacques Vallee. The dynamics of long-term growth. *Technological forecasting and social change*, 7:285–300, 1975.

L. Nottale, J. Chaline, and P. Grou. On the fractal structure of evolutionary trees. In G. Losa, D. Merlini, T. Nonnenmacher, and E. Weibel, editors, *"Fractals in Biology and Medicine", Vol III, Proceedings of Fractal 2000 Third International Symposium*, pages 247–. Birckhäuser Verlag, 2001.

Anders Sandberg and Nick Bostrom. Whole brain emulation: A roadmap. Technical Report #2008-3, Future of Humanity Institute, Oxford University, 2008. http://www.fhi.ox.ac.uk/reports/2008-3.pdf.

Juergen Schmidhuber. New millennium ai and the convergence of history. *Challenges to Computational Intelligence*, 2006.

John Smart. Evo devo universe? a framework for speculations on cosmic culture. In Steven J. Dick, editor, *Cosmos and Culture*. NASA Press, forthcoming.

Ray J. Solomonoff. The time scale of artificial intelligence: reflections on social effects. *Nort-Holland Human Systems Management*, 5:149–153, 1985.

John Maynard Smith and Eörs Szathmáry. *The Major Transitions in Evolution*. Oxford University Press, 1995.

H. Saleur and D. Sornette. Complex exponents and log-periodic corrections in frustrated systems. *J. Phys. I France*, 6(3):327–355, mar 1996.

R Taagepera. People, skills, and resources: an interaction model for world population growth. *Technological forecasting and social change*, 13:13–30, 1979.

Valentin Turchin and Cliff Joslyn. The cybernetic manifesto. http://pespmc1.vub.ac.be/MANIFESTO.html, 1989.

Valentin Turchin. *The Phenomenon of Science. A cybernetic approach to human evolution*. Columbia University Press, New York, 1977.

Time Tyler. The singularity is nonsense. http://alife.co.uk/essays/the_singularity_is_nonsense/.

Vernor Vinge. The coming technological singularity: How to survive in the post-human era. Number NASA CP-10129, 1993.

Heinz von Foerster, Patricia M. Mora, and Lawrence W. Amiot. Doomsday: Friday 13 november, a.d. 2026. *Science*, 132:1291–1295, 1960.

Roland Wagner-Döbler. Rescher's principle of decreasing marginal returns of scientific research. *Scientometrics*, 50(3), 2001.

Eliezer S. Yudkowsky. Three major singularity schools. `http://yudkowsky.net/singularity/schools`, 2007.