Qeios

The Split Singularity Hypothesis

Deep Bhattacharjee, Sanjeevan Singha Roy, and Riddhima Sadhu

Preprint v1

June 30, 2023

https://doi.org/10.32388/6K0DYU

The Split Singularity Hypothesis

Deep Bhattacharjee[1+]
Sanjeevan Singha Roy[12]
Riddhima Sadhu[2]

[1]Electro–Gravitational Space Propulsion Laboratory, India
[2]Birla Institute of Technology, Mesra, Ranchi, 835215, India

Abstract: *Based on the hypothesis of the evolving machine intelligence that in essence gives birth to super-intelligent machines, this can be said that as proposed in the previous theories there will be an intelligence evolution hitting a singularity. In this paper, we have hypothesized that machines may split into two categories based on the 'type' of intelligence explosion resulting in one 'type' hitting hard singularity while one 'type' hitting soft singularity for a specific blow-up case called errored singularity marking the endpoint is discussed.*

Keywords: *Soft Singularity – Hard Singularity – Errored Singularity*

[+]Corresponding author email: itsdeep@live.com
[12]Authors have no conflicting interests related to this paper.

INTRODUCTION

Intelligent machines will rule the future world is a popular saying where starting from the initial hypothesis of Nick Bostrom, several theories have been proposed to depict the danger posed by artificially intelligent machines. This fear of danger leads to several nodes of this theory where terms like seed singularity, intelligence explosion, super-intelligent machines, and machines capable of self-reproduction and all of them pointed in a single way that is towards the harm or posing danger to humanity. Many scientists and physicists have come forward to discuss this issue and the evolution has been marked with the machine via the civilization scales where in future there will be machine domination leading to the making of popular movies and several forms of digital arts regarding how those machines might look into future. This paper will address the issue mathematically through a channelized flow where the transition is shown with the inducement of splitting the singularity into three forms: soft singularity, hard singularity, and errored singularity. This segregation is essential for the demarcation of a specific point where the friendly AI machines will be channelized towards non-(friendly AI) machines which will again roll over via a $\eta \times \eta$ matrix that shows the entire signature of evolution being the operator coefficient of the matrix order[1-3].

PART – I

The linearity of time is the factor for progression and that progress can be stated as evolution where there exists some specific criterion relating all the three factors, the progression, temporal evolution and the elimination of infinity related to that evolution as in any case if the specific domain of progression limits to infinity then it is something that is not capable of human minds to imagine and not even reached in any form of civilization scales as depicted by Sagan and Kardashev.

Therefore, if the spacetime progression can be termed as $\{(\sigma, \rho), \nearrow\}$ then this can be deduced as the temporal progression ρ which is indeed evolution; is a dependable factor on the flow of time $\nearrow$ which must be less than infinity when we split space into several segregations such that each segregation marks each point of evolution such as,

For the segregation parameter of space σ: one can easily split the space via k total units with a constant time ρ to depict as[4-7],

$$\rho \sum_{\mathrm{k}=1}^{\mathrm{n}} \sigma_{\mathrm{k}} \;\exists n \neq \infty$$

Where the progression can be measured in the following way through the setting up of limits,

$$\nearrow \cong \lim_{\rho \neq \infty} \left(\rho \int_{\sigma_{\mathrm{k}}} \Lambda \right)^{(0)}$$

$$\nearrow \cong \lim_{\rho \neq \infty} \left(\rho \int_{\sigma_{\mathrm{k}}} \Lambda \right)^{(1)}$$

Thus, it is now easy to represent the evolution parameter without taking such equations through $(0,1)$ norms depending on the evolution factor Λ in three categories where there exist three projections from the present spacetime: $\{(\sigma, \rho)_{\mathrm{P}}, \nearrow\}$ taking the evolution pattern as,

Iff $\{(\sigma, \rho)_{\mathrm{P}}, \nearrow\} \cong \zeta$ then,

$$\begin{aligned} \partial^{(0,1)} &:= \zeta \hookrightarrow \Lambda^{(0,1)}\big|_{\mathrm{c}} \\ \partial^{(0)} &:= \zeta \hookrightarrow \Lambda^{(0)}\big|_{d} \\ \partial^{(1)} &:= \zeta \hookrightarrow \Lambda^{(0)}\big|_{\mathrm{e}} \end{aligned}$$

For a scheme to be determined $\mathcal{E}_{\mathrm{c}|_{\mathrm{de}}}$ that acts on a basis of probability function $\mathcal{O}$ to exist in the above-mentioned parameter Λ as the singularity parameter which in the way takes three values for the scheme $\mathcal{E}_{\mathrm{c}|_{\mathrm{de}}}$ depending on the dominancy above the probability represented as $\mathcal{O}_{\mathcal{D}}$ which when acting on the singularity factor with the covariant norms in Λ provides the three categories,

$$\sum\nolimits_{\epsilon}^{\times} \left(\bigcup_{\{(0,1)(0)(1)\} \subset \mathcal{D}} \mathcal{O}_{\mathcal{D}} \right)_{\exists c, d, e \; \in \epsilon}$$

Where there lies the generator ϵ_{i} which is a summation of every functional generator that generates the split of the coherent scheme $\left[\left(\mathcal{E}_{\mathrm{c}|_{\mathrm{de}}} \right) \nearrow \right]$ to represent the dominance of probability in the norms $(0,1)$ of the singularity $\wedge$ in a way to encode the evolution parameter $\nearrow$ to represent a closed interval through a finite split-wise basis as,

$$\sum_{\epsilon_{\mathrm{i}}} \times \left(\bigcup_{\{(0,1)(0)(1)\} \subset \mathcal{D}} \mathcal{O}_{\mathcal{D}} \right) \xrightarrow{\Rightarrow} [(\mathcal{E}_{\mathrm{c}|_{\mathrm{de}}}) \nearrow]$$

$$\hookrightarrow \mathrm{c}|_{\mathrm{d}} \xrightarrow{\mathcal{O}_{\mathcal{D}(0)}} \Lambda^{(0)}$$

$$\hookrightarrow \mathrm{c}|_{\mathrm{e}} \xrightarrow{\mathcal{O}_{\mathcal{D}(0)}} \Lambda^{(1)}$$

$$\hookrightarrow \mathrm{c}|_{d,e} \xrightarrow{\mathcal{O}_{\mathcal{D}(0,1)}} \Lambda^{(0,1)}$$

Here, the split can be declared by the vertical bar '|' where we can see three forms of splitting that take place yielding three kings of singularity to say the dominancy in the probability factor $\mathcal{O}_{\mathcal{D}}$ is responsible for the split such that if in the below equation, we denote $\upharpoonleft_0$ then the dominancy factor of (0) is more in $(0,1)$ leading to a hard singularity $\Lambda^{(0)}$ while if we denote $\upharpoonleft_1$ then the dominancy factor of (1) is more in $(0,1)$ leading to a soft singularity and the last of all if we denote $\upharpoonleft_0\upharpoonleft_1$ then this can be easily assumed that the dominancy factor is the same for both the probabilities in a way $\mathcal{O}_{\mathcal{D}}|_{\upharpoonleft_0} \approx \mathcal{O}_{\mathcal{D}}|_{\upharpoonleft_1}$ and we have a null singularity where the singularity won't reach as both the soft and hard is dominant. Now, it takes a bit of explaining about the soft and hard singularity.

A soft singularity is a singularity where these three statements won't fall as,

$$\begin{cases} \text{AI is becoming harmful} \\ \text{AI makes the destruction of humkind} \\ \text{AI will cage humans} \end{cases}$$

The third statement required a bit of explanation as: humans have now caged AI although humans have evolved from AI; in the same context, AI can cage humans in future if they become super-intelligent although they have originated from human minds.

Thus, any singularity that is not soft will be a hard singularity where the intelligent explosion will occur, and super-intelligent machines will rule this planet being capable of self-reproduction and the harmful point of human civilization will arrive.

Thus, to prevent the AI from hitting a hard singularity, the AI can be made to bypass a series of soft singularities where if all of the AI machine gets merged with the soft singularity then there is no point in developing a hard singularity but if a proportion of AI machine gets attached to the soft singularity which means they can't bypass them then they will not anymore be harmful to the humankind, while if any AI can bypass the soft singularity then they get harmful AI: where there can be another scenario which can be merged with the above alternatives as some machine bypass the soft singularity while some can't and if that happens then this is to be determined the extent or magnitude of influence I the machine can provide where the equation can be represented as,

$$I^{\log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right)}$$

Where $\wedge^{(1)}$ is the number of soft singularities and $\wedge^{(0)}$ is the number of hard singularities denoted in natural logarithm in such a case where the number of soft singularities is always more than the number of hard singularities which we will see below (why?) and the magnitude of influence I can take two values where one is forbidden and other should be taken as,

$$\begin{array}{ccc} \text{forbidden value} & \Rightarrow & (-1,0) \\ \text{realistic value} & \Rightarrow & (0,1) \\ \text{fundamental value} & \Rightarrow & (-1,1) \end{array}$$

If soft singularity dominates then the value should be forbidden as between -1 and 0 while if hard singularity dominates then the value should be positive between 0 and $+1$ where a case may arise when both are equal and in this case, the value should be between -1 and $+1$ as the fundamental value where in this case the number of soft and hard singularity will be equal leading to the conclusion,

$$\log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right) = 1$$

Thus, the formula should be modified to only,

$$I$$

Where the calculation of hitting the singularity can be determined on the aspects of the multiplier of the determined computing power $\sqcap$ to attain singularity $\wedge$ with the Sagan–Kardashev Scale: all divided by $I^{log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right)}$ as,

Sagan Kardashev Scale $K = \frac{log_{10} P-6}{10}$ where K is the Kardashev's civilization rating and P is the power consumption.

For $c|_d \equiv c|_e$ the applicable formula will be,

$$\frac{K \times \sqcap}{I^{\log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right)}}$$

For $c|_{d,e}$ the applicable formula will be,

$$\frac{K \times \sqcap}{I} \approx 0 \ \exists \log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right) = \text{undefined iff } \wedge^{(1)} = \wedge^{(0)}$$

PART – II

There has been a concept called friendly AI which supports the formalism of soft singularity $\wedge^{(1)}$ where each machine can be taken as ℓ with a coherent form of all the machines that supports the soft singularity via,

$$\sum_{\mathrm{i}=1}^{>\infty} \ell_{\mathrm{i}}$$

Where the same concept can be used to treat hard singularity $\wedge^{(0)}$ which would be accompanied by machines that are undergoing through a super-intelligent machine for intelligence explosion that can be treated for the specific formalism same as before where a single machine can be taken as O having a coherent form[8-10],

$$\sum_{\mathrm{i}=1}^{>\infty} \mathrm{O}_{\mathrm{i}}$$

Here this is useful to note that soft singularity-obeying machines will always be greater than hard singularity-obeying machines which in turn expresses a positive form of the equation to be valid as,

$$\sum_{\mathrm{i}=1}^{>\infty} \ell_{\mathrm{i}} \gg \sum_{\mathrm{i}=1}^{>\infty} \mathrm{O}_{\mathrm{i}} \xrightarrow{\text{yields positive result for}} \frac{K \times \sqcap}{I^{\log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right)}}$$

Where in the case which is not at all suitable unless the ‘Errored singularity’ is not reached which in turn causes a blow-up of the equation for a negative logarithm giving it undefined where the case can be specifically parameterized by $\overline{\overline{\nabla}}$ which will only occur for,

$$\overline{\overline{\nabla}} \Rightarrow \sum_{\mathrm{i}=1}^{>\infty} \ell_{\mathrm{i}} \ll \sum_{\mathrm{i}=1}^{>\infty} \mathrm{O}_{\mathrm{i}} \xrightarrow{\text{blows up}} \frac{K \times \sqcap}{I^{\log_{10}\left(\wedge^{(1)}-\wedge^{(0)}\right)}}$$

This is determined via a matrix $\eta \times \eta$ with a transition flow marking a boundary of a $\varsigma - parametric$ solution which indeed marked the flow channel of machine intelligence as,

$$\langle \ell_{\mathrm{i}}|_{\eta\times\eta}\rangle \xrightarrow{\quad\varsigma\quad} \langle \mathrm{O}_{\mathrm{i}}|_{\eta\times\eta}\rangle \xrightarrow{\quad\varsigma\quad}]\overline{\overline{\nabla}}]$$

$$\exists \ell_{\mathrm{i}}|_{\eta\times\eta} \neq \mathrm{O}_{\mathrm{i}}|_{\eta\times\eta}$$

Thus, all machines that are there as AI friendly can turn up for a specific case to be a Not-(AI friendly) which via the flow of machine intelligence ς will hit the hard singularity with the extreme boundary being $\overline{\overline{\nabla}}$ where the flow will reverse back in undefined form having a similar number of machines as considered in the $\eta \times \eta$ matrix for a value to be shown with a null value θ representing empty sequence for the machine of another singularity,

$$\begin{bmatrix} \ell_1\theta_1 & \ell_1\theta_2 & \dots & \ell_1\theta_n \\ \ell_2\theta_1 & \ell_2\theta_2 & \dots & \ell_2\theta_n \\ \vdots & \vdots & \dots & \vdots \\ \ell_m\theta_1 & \ell_m\theta_2 & \dots & \ell_m\theta_n \end{bmatrix}_{\equiv \eta\times\eta}$$

$$\downarrow$$
$$\downarrow_{\varsigma}$$
$$\downarrow$$

$$\xrightarrow{\varsigma} \begin{bmatrix} O_1\theta_1 & O_1\theta_2 & \dots & O_1\theta_n \\ O_2\theta_1 & O_2\theta_2 & \dots & O_2\theta_n \\ \vdots & \vdots & \dots & \vdots \\ O_m\theta_1 & O_m\theta_2 & \dots & O_m\theta_n \end{bmatrix}_{>\eta\times\eta} \quad \xrightarrow{\varsigma} \left] \begin{bmatrix} \ell_1 O_1 & \ell_1 O_2 & \dots & \ell_1 O_n \\ \ell_2 O_1 & \ell_2 O_2 & \dots & \ell_2 O_n \\ \vdots & \vdots & \dots & \vdots \\ \ell_m O_1 & \ell_m O_2 & \dots & \ell_m O_n \end{bmatrix}_{\equiv\eta\times\eta} \quad \overline{\overline{\nabla}} \right]$$

Thus the $]\overline{\overline{\nabla}}]$ can be easily determined via a single equation that marks the intelligence explosion for $\equiv \eta \times \eta$ matrix through a closure,

$$\left[\equiv \eta \times \eta_{\ell_i} \bigcup_{\Lambda^{(1)} \xrightarrow{\varsigma} \Lambda^{(0)}}^{\overline{\overline{\nabla}}} > \eta \times \eta_{O_i} \left(\coprod_{\Lambda^{(1)} \equiv \Lambda^{(0)} \subset]\overline{\overline{\nabla}}]} \equiv \eta \times \eta_{]\overline{\overline{\nabla}}]} \right) / \sim \right]$$

CONCLUSION

Three possible scenarios with three types of singularity have been covered in this paper. It has been shown that the ‘soft singularity’ is the ideal AI-friendly machine, which generates a much smaller number of machines that possess super intelligent powers to hit technological singularity or ‘hard singularity’. The channelised flow proposed for the purpose of this paper leads to a point where the lesser components of super-intelligent machines, which have reached ‘hard singularity’, revert back to an equal number of machines present in the initial stage of ‘soft singularity’. At this stage, it can be safely concluded that all machines have undergone an intelligent explosion and have reached the peak of machine intelligence.

REFERENCE

1. Bostrom, N. (2016). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, USA.
2. Russell, S. J., & Norvig, P. (2015). *Artificial Intelligence: A Modern Approach*.
3. Bhattacharjee, D., & Roy, S. S. (2021). Uncontrolled mutation of super-intelligent machines may lead to the destruction of humanity. *TechRxiv*. https://doi.org/10.36227/techrxiv.15081645.v1
4. *What is technological singularity?: AI terms explained - AI For Anyone*. (n.d.). https://www.aiforanyone.org/glossary/technological-singularity
5. Yampolskiy, R. V. (2015). From Seed AI to Technological Singularity via Recursively Self-Improving Software. *arXiv:1502.06512v1 [cs.AI]*. https://doi.org/10.48550/arXiv.1502.06512
6. Bhattacharjee, D. (2021). The Clampdown Effect: On The Expulsion of Super-Intelligence. *Preprints*. https://doi.org/10.20944/preprints202104.0353.v1
7. Boden, M. A. (2018). 7. The Singularity. In *Oxford University Press eBooks*. https://doi.org/10.1093/actrade/9780199602919.003.0007
8. *Ethical Issues In Advanced Artificial Intelligence*. (n.d.). https://nickbostrom.com/ethics/ai
9. Yampolskiy, R. V. (2017). *Artificial Superintelligence: A Futuristic Approach*. Chapman & Hall/CRC.
10. Bhattacharjee, D., & Samal, P. (2023). Erwin Schrodinger and AI. *Authorea (Authorea)*. https://doi.org/10.22541/au.168062605.55771270/v1