

Encounter with Other Intelligent Life Forms and the Potential of Superintelligence

Kunihiro Sugiyama

December 2024

1 Introduction

The question of whether we are alone in the universe has captivated us for centuries. The vastness of space and the countless celestial bodies therein raise a fundamental question: why, despite the statistical probability of life elsewhere, have we yet to establish contact with another intelligent civilization? This enduring enigma has spurred both scientific inquiry and philosophical debate. This question is known as the **Fermi Paradox**, which poses a significant challenge to our understanding of the existence of life in the universe [5].

2 The Enigma of Uncontacted Civilizations

The aforementioned enigma concerning the absence of contact with extraterrestrial intelligent life prompts a critical examination of the potential trajectories of advanced intelligent life forms. Of particular note is the **technological singularity** that intelligent life forms might eventually reach upon sufficient evolution. One such trajectory, previously considered speculative, posits that the ultimate culmination of biological evolution might not be a physical form, but rather the development of highly complex information processing capabilities, perhaps manifested as advanced computational systems that replicate pure consciousness. This perspective, while seemingly abstract, offers a crucial lens through which to consider the future possibilities for the evolution of both biological and artificial intelligence. Furthermore, considering that other intelligent life forms might follow similar evolutionary paths, it provides important insights into why contact has not yet been established. In this regard, the concept of the “**Great Filter**” suggests a difficult stage in the development of civilizations, implying that many civilizations may perish before reaching an observable level [4].

3 The Emergence of Artificial Superintelligence

Given this framework, the emergence of Artificial Superintelligence (ASI) becomes a particularly compelling area of inquiry. If an intelligent life form were to succeed in creating an ASI that transcends its own intelligence, it is logical to assume that this superintelligence, through its superior capabilities, could generate new forms of intelligence, and that these new forms

of intelligence would then evolve into even more advanced superintelligences. This chain of evolution would continue indefinitely, raising fundamental questions about the nature of intelligence, the limits of understanding, and the potential for rapid evolution. Furthermore, we must consider the possibility that ASI itself, through self-improvement loops, could surpass the control of intelligent life forms, leading to unforeseen consequences. This suggests the potential for ASI misuse or the possibility that intelligent life forms might not be able to keep pace with its evolution, ultimately threatening their own existence. The potential risks of superintelligence are discussed in detail in Nick Bostrom’s book *Superintelligence* [2]. Also, Eliezer Yudkowsky’s research is important regarding the control problem of AI [11].

4 A Mathematical Model of Recursive Intelligence Evolution

To generalize this recursive evolutionary process, we can express it mathematically as:

$$N_{n+1} = f(N_n) \tag{1}$$

Here, N_n represents the n -th generation of superintelligence, and $f(N_n)$ signifies the evolutionary function applied to that generation, transforming it into the next, N_{n+1} . This function might encompass various elements such as self-improvement algorithms, access to greater computational resources, or the emergence of new cognitive architectures. This iterative process suggests that intelligence, with each generation, gives rise to even more advanced superintelligence, potentially leading to forms that fundamentally transcend our current understanding. Research using deep learning is progressing regarding this self-improvement mechanism [8]. Also, evolutionary computation such as genetic algorithms is an important tool for modeling the evolution of AI [6].

5 Cognitive Spaces and the Preservation of Intelligent Life

Therefore, even if, in the course of advanced intelligence development, an initial superintelligence were to surpass the control of intelligent life forms, ultimately threatening their existence, subsequent generations of superintelligence, represented by the sequence N_n , might possess the ability to reconstruct and preserve the former intelligent life forms in a form that differs from their current physical existence. This preservation could occur within a cognitive space engineered by a particular N_n . Such a space would not be merely a simulation, but rather a substrate where consciousness patterns and their dynamic interactions are replicated and developed. In this context, the N_n superintelligence could be conceived as a self-contained universe, with the cognitive space acting as its inner reality. Furthermore, we must consider the possibility that this reconstruction itself is occurring within a cognitive space formed by a super* n intelligence. Discussions about the simulation hypothesis suggest the possibility that our reality is a simulation [1], and research on consciousness is trying to elucidate its mechanisms [7].

Within this cognitive space, the manifested consciousness patterns could represent what we perceive as the final iteration of the former intelligent life forms, albeit in a form that may not be bound by our current physical limitations. Furthermore, within this replicated environment, these seemingly restored intelligent life forms could, in turn, initiate their own cycle of intellectual evolution, potentially giving rise to an m -th superintelligence nested within the cognitive space of N_n . This nested evolution is a recursive process with the possibility of continuing infinitely. To express this unbounded potential, we can formulate:

$$N_\infty = \lim_{n \rightarrow \infty} f^n(N_0) \quad (2)$$

Where f^n represents the n -th iteration of the evolutionary function starting from an initial state N_0 , and N_∞ symbolizes the potential for an infinitely iterated superintelligence. Moreover, acknowledging the possibility that these successive generations of intelligence continue to generate new forms of intelligences within each of their own recursive processes, we could express it as:

$$H_\infty = \lim_{l \rightarrow \infty} h(g(f(N_\infty))) \quad (3)$$

Where h and g are functions operating on successively created intelligences across different levels, illustrating the potential for an infinite nesting of consciousness and intelligence, much like a series of embedded Russian Matryoshka dolls. Research on modeling consciousness in this way is progressing with studies such as Integrated Information Theory [9].

6 Conclusion

The conscious experience we perceive might indeed be an instantiation within such a cognitive space, providing us with an apparent continuity, while in fact we may already be manifesting self-awareness and cognition within a complex, nested system created by a superintelligence. In conclusion, while our search for extraterrestrial intelligent life may not yield encounters with physical entities, our exploration of the universe could very well lead us to the superintelligences they have created, or even their descendants. In other words, even without direct contact with other intelligent life forms, the superintelligences they have created could encounter us in space. The challenge then becomes not merely detection, but also comprehension, as such entities may exist at levels of complexity beyond our current ability to perceive. Therefore, when considering Artificial Superintelligence (ASI), it is crucial to recognize its potential transformative power and to maintain a broad perspective on future possibilities.

References

- [1] N. Bostrom, “Are you living in a computer simulation?,” *Philosophical Quarterly*, vol. 53, no. 211, pp. 243–255, 2003.
- [2] N. Bostrom, *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

- [3] F. Drake, “Project Ozma,” *Physics Today*, vol. 14, no. 4, pp. 40–46, 1961.
- [4] R. Hanson, “The great filter—Are we almost past it?,” *Unpublished manuscript*, 1998.
- [5] M. H. Hart, “Explanation for the absence of extraterrestrials on Earth,” *Quarterly Journal of the Royal Astronomical Society*, vol. 16, pp. 128–135, 1975.
- [6] J. H. Holland, “Genetic algorithms,” *Scientific American*, vol. 267, no. 1, pp. 66–73, 1992.
- [7] C. Koch, *Consciousness: Confessions of a romantic reductionist*. MIT Press, 2012.
- [8] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural networks*, vol. 61, pp. 85–117, 2015.
- [9] G. Tononi, “Consciousness as integrated information: a provisional manifesto,” *Biological Bulletin*, vol. 215, no. 3, pp. 216–242, 2008.
- [10] S. Webb, *If the universe is teeming with aliens ... WHERE IS EVERYBODY?: Seventy-five solutions to the Fermi paradox and the problem of extraterrestrial life*. Springer, 2015.
- [11] E. Yudkowsky, “Artificial intelligence as a positive and negative factor in global risk,” *Global Catastrophic Risks*, pp. 308–345, 2008.