Think • Vol 23 • No 68 • Autumn 2024, 63–67

DOI: https://doi.org/10.1017/S1477175624000125

Is the Universe a Dark Forest?

*Amar Desai**

Independent Researcher
*Corresponding author. Email: amar.desai@btinternet.com

Keywords: dark forest; aliens; humanity; extinction

Abstract

The Dark Forest Theory of the universe applies a pessimistic view of humanity to alien civilizations. While many look to space with hope and optimism, the Dark Forest Theory suggests that humanity should be fearful and cautious when expanding beyond Earth and that interaction with aliens could spell humanity's doom. This article will briefly examine the theory and present some arguments against the pessimistic view offered by the Dark Forest Theory.

The Dark Forest Theory of the universe applies a pessimistic view of humanity to alien civilizations. While many look to space with hope and optimism, the Dark Forest Theory suggests we should be fearful and cautious as humanity expands beyond Earth, that interaction with aliens could spell doom for humanity. This article will briefly examine the theory and present a couple of reasons why we perhaps shouldn't be so quick to embrace the pessimism of the Dark Forest Theory. Of course, I won't be able to settle any great cosmic debates over a couple of pages! However, I hope this will be an engaging introduction to a fascinating topic.

The Dark Forest

Cixin Liu's *The Dark Forest*, a science fiction novel in which humanity finds itself at war with a hostile alien race, presents the theory through the following analogy.

Consider a hunter who occupies a forest. They are aware that the forest is occupied by other hunters who seek to kill them, as well as other species who pose no threat. Unfortunately, the forest is dark, and it is impossible to distinguish the dangerous hunters from any other being. Anybody the hunter detects, therefore, could potentially kill them. As such, if the hunter wishes to survive, it is best to adopt the following two strategies:

(1) Remain undetected, lest a hostile party find and kill the hunter.
(2) When another species is detected, shoot and kill them. True, they may be harmless, but they may also be a hunter, and so killing ensures survival by eliminating potentially dangerous parties.

Now consider humanity as the hunter, and the dark forest as space, where vast distances over light years create uncertainty, much like the darkness of the forest. Suppose tomorrow we make contact with another species hundreds of light years away, and detect them as benevolent (given vast distances between civilizations, only very basic information can be exchanged). Furthermore, their technology is so basic that even if they wished to pose a threat to humanity, they would not be able to; they are no match for

https://doi.org/10.1017/S1477175624000125 Published online by Cambridge University Press



our nuclear bombs and other weapons of mass destruction.

However, in the time that it would take for us to encounter them, they may become malicious and dangerous to humanity. Suppose it took only a couple of hundred years for us to be able to make contact (and that is surely a very low estimate given the vastness of space); there is no accounting for how they may have evolved socially and technologically in that time. Consider how it took humanity less than 150 years to go from building the first steam train to the bombings of Hiroshima and Nagasaki. Technology and society can transform radically and explosively in short spaces of time, and so the civilization we detect may be radically different from the civilization we eventually encounter. By the time we encounter them, the alien civilization may be inclined towards humanity's destruction, and capable of enacting it.

In order to ensure humanity's survival, the Dark Forest Theory therefore suggests two strategies:

(1) Remain undetected, so that the new civilization does not become aware of humanity's existence (this option is no longer possible given initial contact has been made).

(2) If it were possible, destroy the other civilization so that they can never become a threat.

It may be tempting to reach out to the other civilization, to share our knowledge with them while accessing their knowledge. The potential benefits to science, the arts and other fields of endeavour would be astronomical. However, it is hard to justify the pursuit of these goods if doing so would risk human extinction.

This can be explored in greater depth through game theory, which examines how choices taken by particular parties (in this case human and extraterrestrial civilizations) can interact to produce different outcomes for the involved parties. This is particularly helpful when one party cannot fully understand how another is likely to act, and so must act on the basis of possible courses the other party will take. In this instance, the vast distances between humanity and the extraterrestrials makes understanding what each other are doing very difficult. Humanity cannot simply look at the aliens and conclude they are going to attack, and the aliens operate under the same limitations. Both parties operate in uncertainty produced by the time lag in communication and the possibility of a technological boom rapidly shifting the balance of power between the two civilizations.

https://doi.org/10.1017/S1477175624000125 Published online by Cambridge University Press

Humanity has two options: to reach out to the extraterrestrials to work together, or to attack and destroy the other civilization. Similarly, the extraterrestrials have two options: attack humanity and seek to destroy them, or reach out to collaborate. Of course, they will have a low chance of success due to their technological inferiority, but if they work hard and develop their technology enough, they will have ample time before humanity can reach them, and can hope for a technological boom that culminates in them possessing better technology than humanity. We can understand the four possible scenarios which would result from the different combinations of approaches through the payoff matrix below. Bold text indicates the decisions taken by each civilization, while italicized text explains the outcome for humanity and the extraterrestrials based on the decisions taken by both civilizations.

‘Consider a hunter who occupies a forest. They are aware that the forest is occupied by other hunters who seek to kill them, as well as other species who pose no threat. Unfortunately, the forest is dark, and it is impossible to distinguish the dangerous hunters from any other being.’

If humanity takes survival as its main objective, then it is rational to attack, as survival if they attack is at the very least likely, and potentially certain. Compare this with seeking collaboration, where one of the two possible scenarios (where extraterrestrials attack) involves certain human extinction. This all seems incredibly bleak and makes the prospect of discovering and interacting with other species horrifying. The choice between mass killing and extinction is not enviable.

There is some historical precedence to support the theory. Consider how much blood has been spilled in the course of history when civilizations encounter each other: where there is a technological gap between two civilizations, it has often been the more technologically advanced nation that exploits and subjugates the other.

Liu also argues that this explains why we detect so little life in the universe: when life is detected by another civilization, it is often eliminated. Detection by another species may not indicate the dawning of a new era of hope and progress for humanity, but may instead mean the complete annihilation of human life.

Allow me to darken things a little further. From 1999 to 2003, a series of messages were transmitted via satellite dishes in Ukraine to nine of the nearest stars, in an attempt to contact alien life. Rather than striving to avoid detection, humanity is broadcasting its existence across the

	Extraterrestrials attack	**Extraterrestrials seek collaboration**
Humanity attack	*Likely human survival* *Likely extraterrestrial extinction*	*Certain human survival* *Certain extraterrestrial extinction*
Humanity seek collaboration	*Certain human extinction* *Certain extraterrestrial survival*	*Certain human survival* *Certain extraterrestrial survival*

https://doi.org/10.1017/S1477175624000125 Published online by Cambridge University Press

cosmos. If Dark Forest Theory is to be believed, we must hope these signals are never detected by alien life. Human designs upon space will only make us more detectable: today we discuss colonizing Mars, and in future eras we may cast our eyes upon other galaxies. Even if one species detects humanity and chooses to ignore us, perhaps another may detect us and decide to destroy us.

Potential Objections

Perhaps there is hope. There are many objections to the Dark Forest theory, what follows below are just a couple of them:

(1) Dark Forest Theory is flawed because it seeks to predict alien behaviour through human paradigms, using inductive reasoning. Simply because there has been bloodshed between human civilizations in the past, that does not necessitate that relationships between human and alien civilizations would be defined through bloodshed.

Suppose I have never seen a frog (or even a picture of a frog) before, but encounter a blue poison dart frog on my travels through the unexplored Amazon. I spot seventy such frogs during my travels, and they are all blue. From this, I conclude that all frogs are blue. I have reasoned inductively, moving from my specific experiences of frogs to the more general claim that 'all frogs are blue'. Of course, I am wrong, and have made a flawed claim through induction. Not all frogs are blue.

Similarly, positing that interactions between civilizations will be consistent over time may employ falsely inductive reasoning. 'Where human civilizations meet, there has historically been bloodshed, and therefore any human encounter with an extraterrestrial civilization is likely to culminate in bloodshed' may similarly be a falsely inductive claim. Indeed, perhaps trying to understand aliens as humans through approaches such as game theory is futile: alien minds might be so radically different from ours that they are beyond our understanding. Perhaps aliens have different ways of behaving that don't follow the patterns we observe in human history, and were they to detect us, their approach would be peaceful and benevolent.

On an empirical level, one might also argue that meetings between civilizations have not always historically been violent. History has seen war between civilizations, but it has also seen trade, positive cultural exchange and the formation of international organizations. There has been war, but there has been cooperation too. Of course, as an empirical argument with vast scope, this debate cannot be resolved in a brief philosophy article. Perhaps one would not be justified in extrapolating from human history that encounters with other civilizations are likely to be violent. Similarly, one may not be justified in applying human paradigms of thought to model interactions with alien species, as aliens may act in a way we simply cannot predict.

(2) The above application of game theory forces the presumption of uncertainty as an unavoidable fact, but perhaps uncertainty can be mitigated. It is true that humans initially cannot be certain of alien actions due to the vast distances between the two civilizations, and that aliens similarly cannot be certain of human actions. However, this uncertainty can be overcome due to both civilizations developing their communications technology and reaching out to develop a mutual understanding. This increases mutual visibility, and so allows aliens and humanity to build trust, as one civilization could not send a fleet of warships or nuclear warheads to the other without being detected and risking retaliation. As such, it is more rational for both civilizations to use communications technology to eliminate uncertainty, which in turn facilitates collaboration. Through

https://doi.org/10.1017/S1477175624000125 Published online by Cambridge University Press

overcoming uncertainty, the dilemma outlined in the above payoff matrix may be avoided altogether.

‘ It may be tempting to reach out to the other civilization, to share our knowledge with them while accessing their knowledge … However, it is hard to justify the pursuit of these goods if doing so would risk human extinction.’

(3) Even if destroying alien civilizations when they are encountered may be the best way to ensure humanity’s survival, the act of destroying a civilization is so immoral that it is never justified regardless of the consequences. Consider the amount of suffering humanity might cause, or the billions of lives that would be ended. As such, independently of the theory’s factual accuracy, it is not right to respond to Dark Forest Theory by adopting a policy of aggression. This is not a disagreement with Dark Forest Theory itself, but rather a comment on how humanity should respond to the theory on an ethical level, and it remains true that we would be wise not to broadcast our existence to the universe.

Given the grim picture of the universe the Dark Forest Theory paints, we must hope that one of the factual objections to it is to be trusted. The alternative may be abhorrent.

Amar Desai

Amar Desai is an English teacher based in Oaxaca, Mexico. He completed his BA in Philosophy at the University of Birmingham in 2016.

Cite this article: Desai A (2024). Is the Universe a Dark Forest? *Think* **23**, 63–67. https://doi.org/10.1017/S1477175624000125

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Royal Institute of Philosophy

https://doi.org/10.1017/S1477175624000125 Published online by Cambridge University Press