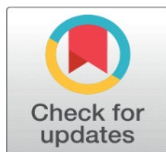
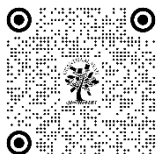


PLAUSIBILITY OF FICTIONAL AI COGNITIONS IN ISAAC ASIMOV'S *I, ROBOT*

Callaghan Gerard Patrick ¹, Karuppasamy Selvam ²

¹ Ph.D. Part-Time Research Scholar of English, Periyar Maniammai Institute of Science & Technology (Deemed to be University), Thanjavur, Tamil Nadu, India

² Associate Professor and Head, Department of Languages, Periyar Maniammai Institute of Science & Technology (Deemed to be University), Thanjavur, Tamil Nadu, India



DOI

[10.29121/shodhkosh.v5.i6.2024.4238](https://doi.org/10.29121/shodhkosh.v5.i6.2024.4238)

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Copyright: © 2024 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



ABSTRACT

Artificial Intelligence has become a vital area to explore in literature as significant importance has been given to the development of technologies, particularly computational technologies and robotics. In contemporary reality, it has become important to understand the standards in which Artificial Intelligence operates and functions in various disciplines that make life more efficient and convenient for human beings. Hence, it is essential to understand the limitations of its functions and operations. Science Fiction is an important genre in English Literature that explores possibilities for Artificial Intelligence development through sub-genres like Robotic Dystopias or Utopias, where realities about uncertain futures are assumed and visualized through the author's imagery and the reader's interpretation. The title "Plausibility of Fictional AI Cognitions in Isaac Asimov's *I, Robot*" is the study of the literary work of art by Isaac Asimov to highlight the interpretation made by the author about a 21st century reality in which almost all areas have robots programmed to perform various tasks. In a reality which is reliant on robots, it is extremely important to understand the ethics that govern them. The robots in the fix-up novel seem to be governed by the fictional: "Three Laws of Robotics". Artificial Intelligence plays a role in the development of every robot's positronic brain. In the series of stories that Isaac Asimov takes the reader through his work; a greater meaning in the metallic creatures is discovered, that the robots are capable of cognitive function that is hard to differentiate between living human beings. This brings an awareness that there is a plausibility in applying the same ethics pragmatically by understanding the ontological significance of the robot characters in the novel. This will provide scope for areas in philosophy, ethics, ontology and AI cognitive studies that are interdisciplinary and not limited to the field of literature.

Keywords: Science Fiction, Robotics, Positronic brain, Artificial Intelligence, Cognition, Robot Ethics

1. INTRODUCTION

Isaac Asimov's *I, Robot* is a fix-up novel. (a fix-up novel is a novel like a book that has a series of stories. Some of them are not entirely connected or events that occur happen at a different time in the textual reality). The novel opens up with a discussion between two scientists who work at U.S. Robots. Susan Calvin had earned her Ph.D. and joined United States Robots as a Robopsychologist. Alfred Lanning had become the director of research for the company. They had both contributed in years of experience for the company on research and development as they watched robots and humans face unprecedented predicaments. The author Asimov uses the character, Susan Calvin to propagate the ideals of his narrative where she recalls the incidents she had heard of or witnessed in person as a flashback story. The series of stories begin to speak about various scenarios where robots conflict with themselves or their existence, robots in a positive or negative relationship with humans; as well as stories of robots that were broken because of "The Three Laws of Robotics" conflicted with each other when a human was subject to harm or was about to be killed. The fictional "Three

laws of Robotics' are a concurrent plausibility that would work in bridging the communication gap of understanding Robotic Artificial Intelligence Cognition. In an interview with Horizon's Past Prediction on 28th July 1965, Asimov said:

The true value of science fiction to me rests in the fact that it permits speculation and makes it respectable, the scientist is always inhibited somewhat when it comes to speculation. He can't or feels he can't put himself in a position where he seems to be going beyond that which the cold logic and evidence makes it reasonable. The science fiction writer can leap across chasms where no evidence has yet filled in matters and speculate in directions which might seem silly perhaps to a solid scientist and yet such speculation is important and is useful today more than ever before, tomorrow more than today; things are moving so quickly that it is unsafe to go forward blindly anymore, one must try to foresee where it is that one is going as much as possible and also foresee what the reactions might be. (BBC Archive 4:10-5:21)

Asimov is one of the greatest pioneers in Science Fiction of his time. He believed in a world where humans and robots co-existed despite their faults and failures. The robots here in his fix-up novel are not destructive by nature and they are only subject to change their behaviour pattern when there is a conflict with 'The Three Laws of Robotics'. Unlike Science, Science Fiction can move across the barriers that limit speculation and looks beyond interpretation. It lays a foundation and creates a stability in bringing awareness to an unforeseen future in an ever-changing world of science and technology.

Research in the sub-genre of Artificial Intelligence in Science Fiction will pave the way for scope in interdisciplinary research in the fields of philosophy, psychology, AI ethics, Ontology and Cognitive Studies. While doing a review of literature, it is found that *I, Robot*, the particular novel by Isaac Asimov have been extensively researched in different interpretations which are concerned mostly with the sub-discipline of social ethics, ontology and cognitive studies. However, the title *Plausibility of Fictional AI Cognitions* is a novel research title chosen to focus on the area of Science Fiction Studies in literature. The title of this paper is deliberately made contradictory to show that the Fictional work of Isaac Asimov can be proven to make sense in today's reality by understanding psychological behavioural patterns in the fictional robot characters. 'Plausibility' according to the Oxford Languages dictionary defines it as the quality of seeming reasonable or probable. Cognition Studies is a broad discipline which has its application in philosophy, sociology, psychology, behavioural science and ethics. 'Cognition' according to the Oxford Languages Dictionary is defined as the mental action or the process of acquiring knowledge and understanding through thought, experience, and the senses. An introduction to the definitions in the title is to show the relevance of exploring new avenues in literary research. Highlighting the narratives that voice the unique robot characters would display their cognitive ability; an ability which shows signs of higher intelligence, something that can be compared to human intelligence. The only difference is that since robots are created by men, their intelligence is artificial. Focusing on the approach to analyse the fictional work of Isaac Asimov, the most noticeable aspect is that the characters that are highlighted are not humans but robots whose cognitive attributes are acknowledged to get deeper insight on how the author's idea of a future robot society can be made possible in reality as well without the fear of knowing that robots will someday rule us as long as they can perceive, reason, understand, and help us humans not as slaves but as people.

In the first chapter *Robbie*, the readers are introduced to the story of a nursemaid bot that was bought by Mr. George Weston for his daughter Gloria. It comes to notice that Mrs. Grace Weston the mother of Gloria does not approve of the robot and wants her child to have a 'normal' childhood. The little girl loves the robot as a friend and sees it as a person. Mr. Weston believes that he has made a good investment in the robot and sees Robbie as a good commodity and is proud that it has saved him time and money since he can keep Gloria happy all the time. Whereas her mother Mr. Weston calls it a "terrible machine" and fears that Robbie may do something to Gloria. Mr. Weston tries to convince his wife to keep the robot and that it is good for Gloria, he also tries to explain the 'The Three Laws of Robotics' to her to no avail. Mrs. Weston eventually finds a way to get rid of Robbie through Mr. Weston and this has broken Gloria's heart. Mr. Weston describes the purpose of Robbie to Mr. Weston as such: "... He just can't help being faithful and loving and kind. He is a machine made so. That's more than you can say for humans" (Asimov 16). This shows that the artificial intelligence of Robbie has cognitive capabilities more or less like human beings. Robbie is not capable of verbal communication because he is designed that way, but that does not make him incapable of communication. However, he is fluent communicator with Gloria and he is able to express his wants as well. When asked by Gloria that he wants a story he nods his head rapidly and uses non-verbal gestures like drawing a semi-circle in the air with one finger to denote that he wants Gloria to tell him the story of Cinderella. Robbie has the ability to express his desire for something as well. When Gloria had told Robbie about the visvox show he had been excited and was in expectation that Gloria would take him there one day.

When Robbie was sold back to the U. S. Robots and Mechanical Men, Inc. by Mr. Weston there was an incident that was inevitably created which made Gloria come in contact with Robbie in dire circumstances. Gloria had found Robbie at the company and started running in wild excitement without knowing that there was a tractor beside her that could mow her down in an instant. The overseers were humans and Mr. Struthers tried to wave at them to make it stop however their reaction time was slow in comparison to Robbie who acted immediately and with precision to save Gloria. This incident proves that the robot Robbie is only capable of making higher order cognitive functions that are more efficient than human cognition.

In the chapter: *Reason*, Gregory Powell and Mike Donovan are reinstated to the Space Station. The two had just put up a new model Robot QT-1 called Cutie. This robot was a bit different from the previous ones as it showed cognitive behaviour right from the time it was created. Powell had explained to Cutie that it is the first robot that displayed curiosity of its own existence. Cutie was still in disbelief even after Powell had explained to him that he was created by him. He seemed unconvinced and exclaimed: "But where do I come in, Powell? You haven't explained my existence." (Asimov 44). This is because the robot's category perception and pattern recognition were blank since it had no input of what it was and was not able to confirm what it was, despite it being curious about the particular objects. Therefore, it found it hard to accept the facts that were presented by Powell. It also raises questions pertaining to ontology about the nature of being. Cutie finds it hard to accept being created by humans and remains skeptical about it because his creators look nothing like him. "I say this in no spirit of contempt, but look at you! The material you are made of is soft and flabby, lacking endurance and strength, depending for energy upon the inefficient oxidation of organic material- like that." (Asimov 47). Furthermore, Cutie believes that the Energy Converter is his creator because it is stronger than itself. The cognitive skill of deduction is most prominent in the robot Cutie. However, it does not account to anything positive or progressive. Cutie come up with an explanation that the Master had created humans as the lowest type which is most easily formed, then later they were replaced by robots which was the next higher step and finally he had created him to take the place of the last humans. When the robots begin worshiping the Energy converter and saying, "There is no Master but the Master and QT-1 is his prophet!" (Asimov 49) it makes us question whether the robot is able to reason. "he's a reasoning robot - damn it. He believes only reason, and there's one trouble with that-" (Asimov 55). This is true because Powell believed that when he created Cutie, he would be able to reason and make the right judgments serving the purpose of the space station coordinator and in charge of all the other robots in the station. Powell states a relevant question of how they can trust it with the station if it does not believe in Earth. The design of the robots' mindframe is cognitive but its cognitive abilities like deduction and reason functions only within 'The Three Laws of Robotics'. This shows that a reasoning robot is good at doing its job but not good in communicating and accepting things outside its reality.

Upon reading Lee McCauley's article: *AI Armageddon and the Three Laws of Robotics*, where he mentions that Isaac Asimov has seen robots as advanced tools and that humans would certainly create a fear of intelligent robots and has thus implemented the three laws of robotics as a safeguard to protect humans from robots in case of a mishap of circumstance. In expressing the challenges faced with practically implementing the laws, McCauley says that there is ambiguity present in the Three Laws of Robotics; particularly the first law as it deals with robots preventing "harm" to come to human beings. He mentions that software experts have expressed their difficulty in translating software code for robots like the word "harm," "intentions," "self," and "other" as the present computing systems are only capable of dealing with logical and numerological problems. This therefore causes cognitive challenges for robots as they are limited in their ability to identify or recognise a human as an individual and may also find it difficult to differentiate between the meaning of the words in a command and to what extent a particular law needs to be applied. As these laws were created to make sure that robots can understand and practice the ethical guidelines while interacting with humans and other robots. Lee McCauley further discusses about a few Alternative frameworks to the Three Laws of Robotics that would help fill the noticeable gaps that are present in the existing Laws of Asimov. He mentions that a researcher named Clarke has proposed additional changes and modifications to the Three Laws of Robotics so as to be better suited to the latest technological and ethical considerations in Artificial Intelligence and Robotics. Clarke's proposed sub-laws were created to focus on researchers and scientists who develop intelligent machines, making sure to take the Robotist's Oath into account while practicing caution, personal responsibility and ethical conduct.

Science Fiction Literature in contemporary times, especially during the pandemic, has opened out and new emerging trends in science fiction and it has contributed to new literatures. Sub-genres like Artificial Intelligence have found its way into literary research, which is interdisciplinary. Pseudo-science and basic understanding of Cognitive Science and ontology can create a lot of new scope for literature aspirants. Also, having a deeper understanding of Science Fiction

genres and literature can be rewardable in this area of research. Scope for this literature would help people who are in technical fields as well by understanding the philosophy of life and existence of both robots and people. It can also pave the way for ethical philosophy studies. Isaac Asimov's *I, Robot* is essentially a guidebook which shows how people should prepare for a society to live in harmony with robots and not to be in fear of them. In every chapter of the novel, the reader is introduced to stories about Robot- Human relationships which are perfect for analysing the cognitive behaviours of robots. Robbie the nursemaid bot shows significant human-like behaviours when interacting with Gloria, especially when he goes all out to save Gloria at the factory. In the chapter 'Reason' the readers are introduced to the robot QT-1 who displays the cognitive function of being curious, later this function is flawed because of its restriction of 'The Three Laws of Robotics' as its curious behaviour creates a sense of distrust and misunderstanding between the humans. From a companion robot to a robot that is confused because it does not want to harm itself and to a robot that puts reason before everything; every one of the robots has a purpose to serve cognitive functions, but just like humans, they are defined around the purpose and the laws that govern them. Understanding how fictional robots behave under fixed norms would shed light on how a robot or AI would display Cognitive behaviours which would help develop new avenues for research in interdisciplinary areas of humanities and Cognitive Studies.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

- Asimov, Isaac. *I, Robot*. Penguin Random House, 2004.
- Languages, Oxford. *Paperback Oxford English Dictionary*. 7th ed., Oxford University Press, 2013.
- Frana, Philip L., and Michael J. Klein. *Encyclopedia of Artificial Intelligence: The Past, Present, and Future of AI*. Bloomsbury Publishing USA, 2021.
- "1965: ISAAC ASIMOV's 3 laws of ROBOTICS | Horizon | Past Predictions", YouTube, uploaded by BBC Archive, 22 Feb. 2022, <https://www.youtube.com/watch?v=P9b4tg640ys>. Accessed 7 June 2023.
- McCauley, Lee. "AI Armageddon and the Three Laws of Robotics." *Ethics and Information Technology*, vol. 9, no. 2, Aug. 2007, pp. 153–64, <https://doi.org/10.1007/s10676-007-9138-2>. Accessed 10 June 2023.
- Murphy, Robin R., and David D. Woods. "Beyond Asimov: The Three Laws of Responsible Robotics." *Machine Ethics and Robot Ethics*, Routledge, 2020, pp. 405–11, <https://doi.org/10.4324/9781003074991-35>. Accessed 10 June 2023.
- Clarke, Roger. "Asimov's Laws of Robotics: Implications for Information Technology." *Machine Ethics and Robot Ethics*, Routledge, 2020, pp. 43–52, <https://doi.org/10.4324/9781003074991-4>. Accessed 11 June 2023