I,Robot: An Ethical Interrogation of AI

Dr. Raghavendra H M
Associate Professor of English
Government First Grade college, Tumkur-572102, India

Abstract- "I, Robot" is an exceptional work in the genre of science fiction by Isaac Asimov. It explores the robot-human relationship, analysing the ethical and moral implications of artificial intelligence in the future society. Asimov depicts the future of robotics based on the famous "Three Laws of Robotics", which is the basis of perceiving the robot-human relationship in the human dominated society. The present article critically analyses the depiction of AI- Artificial Intelligence and HI- Human Intelligence in a fictional futuristic society as depicted in the stories of " I, Robot" by Isaac Asimov. The themes of AI morality ethics and literature are discussed from the beginning of the last century initiated by Isaac Asimov's, "I, Robot" stories. These stories have contributed greatly to the genre of future science fiction with reference to AI and robotics study.AI backed robots depicted by Asimov are not simple machines which may be built to be human servants. They are conscientious beings like man with the capacity to work on the delicate ethical rules and loss even beyond the restrictions placed by programming inside them. This portrayal of robots by Asimov highlights many invaluable issues for the present day robotics study and innovations in the field of artificial intelligence.

Keywords- Robots, Artificial Intelligence, human, morality, Ethics, human safety.

I. Introduction

Isaacc Asimov's "I, Robot" published in 1950, is an epoch making futuristic science fiction portraying the AI and human relationship. It also analyses the moral, ethical and philosophical enquiries about this subject. "I, Robot" published in 1950 is a bunch of independent short stories connected with each other analysing man - robot interconnectedness. The robots portrayed in these short stories are the most advanced AI supported machines. The core of these stories is "the three laws of robotics" made famous by Asimov. These laws are constructed to safeguard human beings from robotic violence and to protect the society from artificial intelligence.

These stories also defeat the implications of autonomy, morality and ethics concerning artificial intelligence in society. Considering huge progress in development of AI and the ever increasing dependence on machines in the form of robots by human society. The author brings out the themes of man - AI relationship,natural and artificial intelligence, morality and ethics concerning the use of robots and the advantages of AI to man and the serious threats AI may pose to human beings in future. "I, Robot" stories in general discusses the clash of AI with human intelligence it also brings out the most important issues such as the autonomy of robot and the danger it poses to man, the limitless possibilities of utilising AI and robots and its consequences and the moral dilemma it poses and the challenges in front of human society in tackling this danger.

II. Literature Review

The themes of AI morality ethics and literature are discussed from the beginning of the last century initiated by Isaac Asimov's, "I, Robot" stories. These stories have contributed greatly to the genre of future science fiction with reference to AI and robotics study. The first law of his three laws of robotics state that, the robots must at any cost never disobey any order by a human being. The second law states, at any cost, to safeguard the life of a human being and the third law says, robots may protect themselves only if it does not conflict with the first two laws stated above.

"**The Three Laws of Robotics:**
1. : A robot may not injure a human being or, through inaction, allow a human being to come to harm;
2. : A robot must obey the orders given it by human beings except where such orders would conflict with the First Law;
3. : A robot must protect its own existence as long as such protection does not conflict with the First or Second Law;
The Zeroth Law: A robot may not harm humanity, or, by inaction, allow humanity to come to harm."(Asimov, I, Robot (Robot, #0.1)

These Three Laws of robotics are considered step stones for discussion about AI morality, ethics and artificial intelligence in future.
Many scholars like Abraham Paul M, Staurt Renter, Muller, and Ord have noticed the impact of these three laws in designing the AI related ethics for the artificial machines that are robots. These laws presented by works as cornerstone to enquire about the conflict between AI programming and AI autonomy. It also enquires if complete control over AI is needed or not.

The main theme in the stories is robotic autonomy. The stories " run around" "reason " included in the short stories series " "I, Robot"" the robotic machines behave according to the implications of the three laws which leads to ethics which is really unexpected. It presents the ambiguity in these relationships.

The autonomy of artificial intelligence is the basis for all interactions and discourses regarding AI-like decision making systems as in Abei (2016). These decision making systems are programmed to make decisions on the basis of enormous data which may work in so many methods that pose a huge challenge to human expectations and ethical values.

A Robot says, " I, on the other hand, am a finished product. I absorb electrical energy directly and utilize it with an almost one hundred percent efficiency. I am composed of strong metal, am continuously conscious, and can stand extremes of environment easily. These are facts which, with the self-evident proposition that no being can create another being superior to itself, smashes your silly hypothesis to nothing."(Asimov, I, Robot, Robot, #0.1)

The stories present an inquiry around AI and robotics regarding the capacity of robots in taking independent ethical decisions are very well discussed in detail by Isaac Asimov in these short stories.

These intelligent machines depicted in the series "I, Robot", are never simple machines but they are some kind of intelligent beings with the ability to make their own decisions and choose between what is right and what is wrong considering the framework of the three laws. These discussions pose questions, like- the robots can really follow the moral and ethical rules. It also inquires about the behaviour of society regarding robotic choice making(Carlsmith,2022).

These discussions are very much relevant today as the AI technology is developing in a rapid pace. The predictions of a symbol regarding AI are first in this field of science fiction. This may be fictional but the new growth of AI related technology is bringing a Asimov's predictions into the real world situations. Yudkowsky(2008) thinks, the predictions of highly relevant in the current situation where AI technology is rapidly progressing. So Asimo's stories are highly relevant regarding the protection of human life and society from AI dominance in present and future times.

III. Results

The study of "I, Robot" stories by Isaac Asimov shows numerous findings regarding AI and robotic science.
The AI installed robots behave as if they are giving first priority to the three laws when making decisions and functioning. At the same time these robots also show very sophisticated decision making abilities with reference to ethical and moral laws.
This shows the daunting task in front of the scientists and engineers while designing artificial intelligence technology that can completely comprehend and adapt according to the varied needs of the society and its moral and ethical behaviour.

The three laws of Asimov seem to have a very rigid base. They are not at all flexible to adapt themselves according to their needs. In the stories "I, Robot", AI based robots are forced to confront several moral implications which push them to ethical dilemmas regarding human safety and robot safety. These robots have to understand the AI programming installed in them in a way to decide the two choices between human life versus robot safety.

This inevitability shows that artificial intelligence may face ethically grey areas which pose a challenge to the controlling of robots by human beings.

The "I, Robot" stories prove that AI has the capacity to develop itself like human evolution. It can cross the limits of programming which puts the question of whether artificial intelligence will outsmart human intelligence in the future or not. They can evolve to such an extent that they can be lunatic like a real person, "Sizzling Saturn, we've got a lunatic robot on our hands (Asimov, I, Robot).

On the positive side, AI has the capacity to completely resolve the problem of human error in programming and in general. This prediction by me post a huge apprehension

to contemporary society regarding autonomous decision making capacity of artificial intelligence.

IV. Discussion

The basic concern of Isaac Asimov's short stories collection "I, Robot" is it credits the artificial intelligence's inevitable relationship with man and the society in future. These AI backed robots depicted by Asimov are not simple machines which may be built to be human servants. They are conscientious beings like man with the capacity to work on the delicate ethical rules and loss even beyond the restrictions placed by programming inside them. This portrayal of robots by Asimov highlights many invaluable issues for the present day robotics study and innovations in the field of artificial intelligence.

These stories highlights the capacity of AI to check independent decisions regarding ethically UN clear matters beyond their scripted programs. This tendency is especially mentioned in the story "little lost robot". In this story a robot understands the three laws, but it comprehends them in completely different ways which result in loss of human life. This behaviour is a very crucial matter in understanding AI intelligence today, as the AI machines are used extensively in a very delicate areas such as medicine driver less vehicles criminology forensic science and justice. Today AI program med robots are extensively used in heart surgery, neuro, eye and Orthopaedic surgery.

V. Conclusion

The short stories of a robot by Asimo help us understand the complex behaviour of AI and its influence on human life and society. By portraying the robots tied by the three laws of robotics Asimov presents AI as a boon and a curse to humanity. These stories may be just fictitious representations of AI and robots but they are futuristic in research in the field of Artificial Intelligence and programming robots. AI is the future, it is development and evolving into a huge potential without which human beings may not function even in their day to day activities. So these stories function as a beacon of light to the ship of artificial intelligence research. It gives invaluable guidance in developing smart machines in the future.

Reference

1. Abei, D. (2016). Problems in AI Safety. AI and Computer Technology, 16(6).
2. Abraham, Paul M and Staurt Renter. "Tik Tok and the Three Laws of Robotics." Science Fiction Studies 5.1 (1978): 67-80.
3. Akhmedov, R. (2024). The Theme of AI Technology in American Science Fiction. Tadqiqotlar, 35(3), 84-87. URL: https://tadqiqotlar.uz/new/article/view/2533/2364
4. Asimov, Issac. Robot Visions. New York: Roc-Penguin, 1991. Print.
5. ---. Robots of Dawn. New York: Bantam, 1994. Print.
6. ---. The Rest of the Robots. London: Harper Collins, 1997. Print.
7. ---. Robot Dreams. London: Gollancz, 2004. Print.
8. ---. I, Robot. London: Harper Collins, 2013. Print.

9. ---. Robots and Empire. London: Harper Collins, 1996. Print.
10. Carlsmith, J. (2022). Is Power-Seeking AI an Existential Risk? URL: https://arxiv.org/abs/2206.13353
11. Girisha, D. (2021) Synchronization of Myth and Modernity in the Selected Novels of Isaac Asimov. Annamalai University.
12. Müller, V. (2020). Ethics of Artificial Intelligence and Robotics. Stanford Encyclopedia of Philosophy. Stanford: Stanford University Press.
13. Ord, T. (2020). The Precipice: Existential Risk and the Future of Humanity. London: Bloomsbury Publishing.
14. Yudkowsky, E. (2008). Artificial Intelligence as a Positive and Negative Factor in Global Risk. In N. Bostrom & M. Ćirković (Eds.) Global Catastrophic Risks (pp.308-345). New York: Oxford University Press.