

# Morality of Robots

<sup>1</sup>Rosy Guha, <sup>2</sup>Dr. Vishwanath Y

<sup>1</sup>Student, <sup>2</sup>Associate Professor

<sup>1</sup>Information Science and Engineering,

<sup>1</sup>New Horizon College of Engineering, Bangalore, India

**Abstract :** Robots have become a part of our daily work environment for the past few years. They are automating a wide range of daily activities in various aspects. Robotic applications have increasingly developed in health care industry, search and rescue operations, tutor our children, self-driving cars, care for our elderly, military robots in automated warfare, to provide companionship etc. Thus, robots are making our lives faster and easier. Nowadays, the potential for human-robot interaction is emerging. As these grow up to be more capable, it is obvious that they will throw an impact ethically, physically as well as emotionally on humans. They will be involved in situations where moral character will be integrated to their interactions with other sentient animals. Robots have to get adapted to the individual to make sure that it is assisting the human morally to respect his emotions, beliefs, and dignity. This paper gives an overview on why morality of machines has become an important aspect for human beings.

**IndexTerms - moral agents, socially interactive robots, artificial agents, human-robot interaction.**

## I. LAWS OF ROBOTICS

In 1947, science fiction author Isaac Asimov introduced his Three Laws of Robotics in a short story collection called *I, Robot*. It has a simple set of guidelines for good robot behavior. The three laws of robotics are: 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 Laws. These laws were created to make the robots safe for human beings.

But Asimov's robots adhere strictly to the laws which may lead to unresolvable moral dilemmas in many situations. If we consider the second law, it will positively conflict if a human asks the robot to protect him from another person, intending to kill the first person. In this case, which side the robot should take if both the persons order the robot to kill the other? The dilemma can result into two possibilities: Action: People get hurt or Inaction: Person gets hurt. Will the robot have loyalty if it is an army commander? This conundrum eventually makes the robot insane.

In Asimov's short story called "Runaround", where two scientists discover that they are running out of fuel. They send a robot named Speedy to collect fuel. But after few hours, they find Speedy running in circles and reciting nonsense. The reason is moral crisis which happened due to the conflict of second and third law of Asimov. He is asked to obey human orders but he is also programmed not to cause himself harm. One of the scientist observes "Rule three drives him back and rule two drives him forward". Thus, it proves that even a transparent, logical set of rules may fail when it is interpreted by the mind of robots.

## II. MORALITY IN ROBOTIC TECHNOLOGIES

In robotics technology, the schematic for moral relationship between the agents is given by:

Programmer-> Robot-> User

To clarify the moral relation between natural and artificial agents, we will discuss the categories of robotic technologies:

### Telerobots

These are remotely controlled machines and do not need complex artificial intelligence to run. They make only minimal autonomous decisions. These robots would be applicable as telerobotic nurses, remotely controlled weapons platforms such as Predator drone and the SWORD, search and rescue operations. The schematic is defined as:

Operator-> Robot-> Victim

The ethical decisions of telerobots is born on designers, programmers, and the users who is operating it. This remote control of robot makes it easier for the operator to make moral decisions.

### Autonomous robot

Autonomy means that the robot is designed to deliver tasks without continuous human intervention, guidance and assistance. Autonomous robots are capable of making at least some of the major decisions using their own programming. This may be interesting

and simple in real applications when robot vacuum makes its own decisions on how to navigate a floor that it is cleaning, or when the robot need to be robust to make complex moral ethical reasoning to decide on how to interact with the patient so that, it lifts the interest of both the machine and the patient in equal manner [3].

### III. ROBOTS AND MORAL AGENCY

There are four different views on whether robots can be moral agents or not. The first view says that robots can only do what they are programmed to do and cannot have the autonomous will to be a moral agent. Then the morality of the situation resides fully with users. The second view says that robots are not yet moral agents but they might become so in future, if the intentionality and responsibility can be satisfied. Another view says that issues such as free will and intentionality are still unresolved, “mind-less morality”, should be adopted for non-human entities. Final view says that only an agent that is operated by strict logical reasons has free will, and therefore robots are true moral agents [2].

In Alan Turing’s famous paper of 1950 he introduced a test called as Turing Test which describes the machine’s ability to demonstrate intelligence. This is opposed by Searle by his famous Chinese Room thought-experiment, which shows that it is possible for a computer to pass the Turing Test without possessing genuine understanding and intelligence. He argued against the possibility of “true artificial

intelligence”. He differentiated between two different hypotheses about AI. When an artificial intelligence system can think and have a mind, it is strong AI, whereas when a system act as if it thinks and has a mind, it is weak AI.

If we say that robot is a moral agent, then what does that imply? Should we punish the robots in the same way as humans, when they are responsible for some bad actions? Then, why should we hold the robot responsible for bad actions when we can just switch it off? With growing technology, more advanced robots come into picture which possess more complex features of philosophical views similar to human. An advanced robot may learn how to program itself with ethical principles. So, there may be ways to punish the robot as it may learn. It is also sometimes argued that a robot can never be responsible since it is always programmed by humans. However, even humans are also programmed in some cases. Scientists are working on the area of artificial emotions such as guilt. If robots become aware of the outcome of their actions, and break rules, then they are in the verge of being a human. Concluding this, if we consider robots as moral agents then it is justified to hold them responsible for their actions.

To evaluate the moral status of robots, three things need to be answered: 1) Is the robot autonomous? If that autonomy is able to fulfil the goal and tasks and is not under the direct control of any other agents, then we say that the robot has effective autonomy. 2) Does the robot behave intentionally? If the complex interaction of the robot’s program and the environment causes the machine to behave that is morally beneficial or dangerous, and the actions are probably deliberate, then that robot is considered as moral agent. 3) Is the robot in the position to be responsible? If the robot is able to perform social role that includes responsibility, to create a belief in human that it has the duty to care for its patients, then we can call it as a moral agent. So, a machine, may be robotic caregivers or robotic nurse, will be a moral agent if it is autonomous, behaves in an intentional way and whose programming is complex enough to understand the role of responsibility in caring a patient [3].

### IV. WHEN ROBOTS NEED TO BE MORAL AGENTS?

When computerized automated technologies first appeared in our lives, we were overwhelmed by their capabilities, but as it kept growing, their moral nature of our interactions with them is concerned. We started looking for more products that simulate the biological entities: robot dogs, service robots, robot assistance, humanoid robots, robot nurse, robot army, social robots and many more. Robotics play a greater role in health care industry, social care, child care, decision making, military, domestic chores etc. They are able to fill the gaps in current telehealthcare and home care/self-care provisions [7].

They also support a number of different and “softer” human-robot interaction tasks (HRI) to improve the medical conditions of patients. HRI can be verbal or gesture based interaction which helps in two ways: First, a robot can conduit for socializing, enabling friends and family to get engaged with the user, or for remote communication with a health professional, making them to reach a greater number of patients. Second, a robot can act as an autonomous entity, creating play scenarios to reduce the effect of disabilities, or improving well-being through entertainment and companionship. This robotic therapy helps in improving the social and communication skills of children having such disabilities. Not only that, HRI improves the quality of life by making the robots play a role of coach, motivating people for various tasks like healthy diet, prevent smoking, exercise, balanced life [5].

Robots are able to serve the purpose of humans by being assistive where they acquire some emotional attachment with humans, thus insisting owners to take care of themselves by resembling various child-like behaviors. Robots also turn out to be a good decision makers in many cases by directing humans to the right action that need to be taken. Besides that, robot’s model of empathy is used in virtual agents, to help in developing long-term relationships with humans. These conveys how moral robots are crucial for humans [1].

### V. MORALITY IN HUMAN ROBOT INTERACTION

From the time robots came into existence, researchers are fascinated to make the interactions between a robot and its environment possible. The form and structure of a robot helps establish social expectations. Socially interactive robots should possess the following characteristics:

Express or perceive emotions

1. Learn/recognize model of other agents
2. Use natural expressions (gestures, gaze...)
3. Exhibit distinct personality and character
4. Understand and communicate with high-level dialogue
5. Establish/maintain social relationships

Interactive robots act as partners, peers, or assistants, which means they need to be flexible and user-friendly to humans. They need to have the ability to recognize social context and support a wide range of users of different genders, different cultural and social backgrounds, different ages etc. As they become a part of person's life, robots need to treat him as a distinct individual. With the approach of robot being "functionally designed", the robot may only need to be outwardly/superficially intelligent, even if the internal design does not need to be understanding as the user is not interested in how the mind really works. Even limited social expressions, recorded or scripted speech is sufficient for human-robot interactions.

It is said, the more a robot can perturb an environment, and be perturbed by it, the more it is embodied. This means that social robots do not necessarily need a physical agent. But physical appearance affects interaction in case of a human. A dog robot will be treated differently than the one which is anthropomorphic. A robot which is socially interacting with humans need to be structurally and functionally similar to a human. Artificial emotions, like fear, anger, happiness, sadness, surprise reflects the intention or internal state of a robot and provides feedback to the user. Even, emotional speech is also implemented in few robot systems. But the facial expression in a robot may not be life-like. These facial expressions and emotions are integrated with body language or non-verbal communication of the robot to make it feel life-like. Besides all these, distinct personality differentiates one robot from the other which would make people to be more willing to interact and establish relationship with it. Five common personality types are: tool-like, pet or creature, cartoon, artificial being, human-like. In addition to this, robot's human-oriented perception

capability helps to track human features, interpret human speech, and recognize facial expressions or gestures [4]. Thus, although it is important to enhance the autonomous capabilities, we must not neglect improving the human-robot relationship so that they can participate with full richness in human environment.

## VI. ETHICAL ISSUES IN MORALITY OF ROBOTS

As robotic technology becomes more ubiquitous, the ethical issues are also arising in various aspects. Robots designed as replacement nurses may feel the patients as objects and can cause a possible loss of personalized care. There will be loss of human contact with elderly which can cause a measurable impact on their health and well-being, thus reinforcing the idea of getting deprived which is unethical. Though, artificial intelligence plays a beneficial role in human's life, it should be taken care that AI technology should not be used to replace those positions that require respect and care. A question arises, can a robot assess the mental and physical conditions of the patient? Who should be accountable if a robot responding to all commands of its owner and something goes wrong [6]? So, a right balance need to be maintained between caring an elderly person and protecting them from vulnerable situations that they might come across as a result.

## VII. CONCLUSION

The ethics of artificial intelligence is the cause of ethics of technology which lead to the evolution of ethical robots and other artificially intelligent beings. This roboethics is a boon to the human society in various aspects. With enhancing technology, and considering all the complexities mentioned in this paper, more new robots will be designed that are socially active, morally acceptable, safer, and overall do better job at accompanying those in need.

## REFERENCES

- [1] J. R. Wilson, "Robot Assistance in Medication Management Tasks" from IEEE papers.
- [2] Solace Shen, "The Curious Case of Human-Robot Morality" from IEEE papers.
- [3] John P. Sullins, "When Is a Robot Moral Agent?" from *International Review of Information Ethics*, Vol 6.
- [4] Terrence Fong, Illah Nourbaksh, Kerstin Dautenhahn, "A survey of socially interactive robots" from *Robotics and Autonomous Systems*.

- [5] Torbjorn S. Dahl, Maged N. Kamel Boulos, “Robots in Health and Social Care: A Complementary Technology to Home Care and Telehealthcare” from *robotics journal*.
- [6] Anne Gerdes, “The Issue of Moral Consideration in Robot Ethics” from *ACM SIGCAS Computer and Society*.
- [7] Bernd Carsten Stahl, Mark Coeckelbergh, “Ethics of healthcare robotics: Towards responsible research and innovation” from *Robotics and Autonomous Systems*.

