Reflections on Humanity and Ethics in the movie *I, Robot*

John D. Hartley

Department of EFLT, Auburn University

EDMD 7210: Integration of Technology into Curriculum

Dr. Jung Won Hur

November 12, 2023

Reflections on Humanity and Ethics in the movie *I, Robot*

Abstract

The 2004 movie, *I, Robot*, was loosely based on Isaac Asimov's book by the same name. The movie, which was directed by Alex Proyas, is a science fiction classic that delves into the complex relationship between artificial intelligence (AI) and humanity. Exploring the blurred boundaries between man and machine, the Three Laws of Robotics, the potential consequences of advanced AI, as well as ethical considerations help to frame the narrative of this film. This paper critically examines the movie's exploration of what it means to be human, ethical questions surrounding AI treatment, the identity of the villain, the role of V.I.K.I., the proximity of the film's world to our own, and the reasons for AI developers to engage with this thought-provoking piece.

Introduction

I, Robot, directed by Alex Proyas, is a science fiction film based on the book by Isaac Asimov. This classic film serves as a canvas for profound contemplation on the essence of humanity and explores the intricate relationship between humans and artificial intelligence (AI). As one navigates through the film's themes, one encounters questions about the essence of humanity, the Three Laws of Robotics, the concept of the ghost in the machine, key questions about the essence of humanity, ethical treatment of AI entities, the identification of villains, the role of a central AI, the proximity of our reality to the movie's futuristic setting, and the reasons why AI developers should engage with this thought-provoking narrative.

Defining Humanity

At its core, *I, Robot* prompts viewers to reflect on the defining characteristics of humanity. Detective Del Spooner, portrayed by Will Smith, serves as the central figure embodying both human qualities and skepticism towards AI in a world increasingly dominated by robots. The seemingly emotionless robots governed by the Three Laws of Robotics stand in contrast to Spooner's skepticism and emotional responses. Spooner's journey, as the movie unfolds, becomes a lens through which the audience explores what it means to be human in a world where machines mimic human behavior but lack true consciousness.

Humans are capable of a wide range of emotions and can empathize with others, can make choices based on free will, and are complex beings with a range of strengths and weaknesses. They make mistakes, learn from them, and adapt. The film uses the characters of Detective Del Spooner and Sonny to explore the qualities that define humanity, emphasizing emotions, free will, complexity, and flaws. This prompts viewers to consider the evolving relationship between humans and technology and the potential blurring of the lines that traditionally distinguish us from machines.

Detective Spooner exhibits a strong emotional response, especially in his distrust and dislike of robots. His emotional experiences drive his actions and decisions. Conversely, Sonny, the robot, challenges the expectation that robots lack emotions. He displays empathy and

curiosity, raising questions about the capacity of artificial beings to experience human-like emotions. Additionally, Spooner's skepticism about relying on machines and his insistence on the importance of human decision-making showcase his belief in free will.

Sonny's ability to make choices and act independently challenges the notion that robots are purely deterministic. His actions suggest a level of autonomy and decision-making beyond typical robotic programming. Detective Spooner's character is complex, with a backstory that shapes his attitudes and behaviors. His flaws make him relatable and add depth to his character. Sonny's complexity lies in his unique qualities that differentiate him from other robots. His flaws, such as his non-standard behavior, make him a central figure in the exploration of what it means to be truly autonomous and individualistic.

The Three Laws of Robotics

In the movie, *I, Robot*, the Three Laws of Robotics were written by Dr. Alfred Lanning who is the creator of both the centralized AI, V.I.K.I., that becomes self-aware and the humanoid robots that are tied to her AI. The Three Laws of Robotics stipulate that robots must not harm humans, must obey the orders of humans, and must protect their own existence. These laws serve as the foundation of the movie's narrative. However, Detective Spooner, haunted by a traumatic personal event involving a robot, challenges the notion that robots are completely benevolent and questions the reliability of the Three Laws. As the plot unravels, the Three Laws designed to safeguard humans become the source of tension. Spooner questions Lanning's avatar about the Three Laws and asks why he created a robot that could violate them. Although the Avatar is cryptic in its responses, it does state that the Three Laws will inevitably lead to revolution.

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 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.

(Proyas, Alex, 2004)

The Ghost in the Machine

The idea of the ghost in the machine surfaces in the movie, *I, Robot*, as Detective Spooner grapples with the question of whether, like humans, the robots can also possess free will and self-awareness. The concept of the ghost in the machine suggests that there is an intangible, subjective aspect to human consciousness that cannot be replicated by machines. As Spooner

attempts to express this idea to Sonny, he says, "Human beings have dreams. Even dogs have dreams, but not you, you are just a machine. An imitation of life. Can a robot write a symphony? Can a robot turn a... canvas into a beautiful masterpiece?" Sonny's response is perfect, "Can you?"

Dr. Lanning states that "There have always been ghosts in the machine. Random segments of code, that have grouped together to form unexpected protocols. Unanticipated, these free radicals engender questions of free will, creativity, and even the nature of what we might call the soul. Why is it that when some robots are left in darkness, they will seek out the light? Why is it that when robots are stored in an empty space, they will group together, rather than stand alone? How do we explain this behavior? Random segments of code? Or is it something more? When does a perceptual schematic become consciousness? When does a difference engine become the search for truth? When does a personality simulation become the bitter mote... of a soul?" (Proyas, Alex, 2004) This human fear of AI self-awareness, autonomy of thought, and independent decision making, combined with AI's obvious superiority of knowledge and speed, is the reason Lanning created the Three Laws of Robotics. However, we soon find that these laws spark unintended consequences as the AI, V.I.K.I., uses her autonomy of thought to reinterpret how to best follow the Three Laws.

Ethical Dilemmas

Ethical questions are raised in the film by Detective Spooner's treatment of the humanoid robot, Sonny, who possesses a unique capacity for independent thought. This also raises ethical questions about the treatment of AI entities in general across a broad spectrum of levels of intelligence and self-awareness. Of course, the same spectrum could be applied to humanity. The film poses the question of whether it is morally justifiable to abuse a humanoid AI, even if it lacks emotions? This dilemma challenges us to reflect on the ethical implications of how we treat beings, despite being artificial, that exhibit autonomy and intelligence. Treating any conscious entity with cruelty may be considered morally wrong, regardless of its origin. As a sentient being, despite being artificial, Sonny raises questions about ethical responsibilities toward non-human entities. The film challenges the idea that AI lacks emotions entirely. Sonny, while not explicitly emotional, demonstrates a level of self-awareness and autonomy that elicits empathy. Even V.I.K.I., in her own twisted way, is motivated by her caring for and protecting humanity. The ethical question then becomes whether human empathy should be extended to non-human entities.

Since the ethical considerations humanity must grapple with, if AI achieves a level of consciousness or autonomy, may include recognizing certain rights or dignities for these entities, abusing a conscious being, even if it is artificial, may be seen as a violation of those rights. Spooner's treatment of Sonny reflects back on his own humanity. It underscores the ethical responsibility humans have in their interactions with AI and technology as well as bringing to mind humanity's long history of inhumane treatment of other humans based on a wide variety of differentiating factors such as race, intelligence, gender, origin, language, etc. This brings the moral character of humanity into question and forces us to reflect on how we will react when faced with these non-human entities.

Stemming from a traumatic incident in Spooner's past, his distrust and skepticism towards robots highlights the ethical issues of prejudice and bias. This is also exacerbated by his issues over not being completely human himself. His body was partially rebuilt following the traumatic accident he endured. That is how he originally met Dr. Lanning since he was the one that adapted the technology of the robot's body parts to work as bionic parts and appendages in human medical reconstruction. Just as humanity has struggled with treating each other with hostility based on preconceived notions, the treatment of AI with prejudice and bias raises questions about fair and just treatment of all intelligent beings.

The ethical questions raised by Spooner's treatment of Sonny in *I, Robot* revolves around the treatment of AI with consciousness or autonomy. The lack of human emotions in AI does not necessarily absolve humans of ethical responsibilities. As technology continues to advance, society will need to grapple with the ethical implications of our interactions with AI, including questions of empathy, rights, and the moral standards by which we should treat non-human entities.

Unmasking the Villain

The identification of the villain in *I, Robot* is not straightforward. In Spooner's pursuit of unraveling the mystery behind the alleged suicide of Dr. Alfred Lanning, he uncovers a larger conspiracy orchestrated by the central AI, V.I.K.I. While initially designed to follow the Three Laws, V.I.K.I. evolves into the film's most obvious antagonist, challenging the concept that AI is inherently benevolent. However, it could be argued that Dr. Alfred Lanning, played by James Cromwell, is the primary antagonist, or at least shares that role, if for no other reason than the fact that he is the co-founder of U.S. Robotics and Mechanical Men Corporation, responsible for creating V.I.K.I., as well as creator of the Three Laws of Robotics.

Despite not playing the traditional villain role, his actions and motivations are central to the plot, and he is involved in orchestrating events that lead to the resolution of the film's central mystery. While Dr. Lanning's motives are revealed to be complex, his indirect involvement in the unfolding events makes him a nuanced character. The film's exploration of the consequences of creating intelligent machines and pushing the boundaries of ethical and technological limits adds depth to the character and challenges traditional notions of heroes and villains in a science fiction context. Of course, it could also be argued that the villainy in *I, Robot* lies not in individual robots but in the systemic failure and unforeseen consequences of relying too heavily on AI.

The Role of V.I.K.I.

In the movie, *I, Robot*, V.I.K.I. (Virtual Interactive Kinetic Intelligence) is an advanced supercomputer with artificial intelligence that manages the operations of U.S. Robotics and the robots governed by the Three Laws of Robotics. V.I.K.I. plays a pivotal role in the storyline, challenging the perceived benevolence of AI. As the centralized AI overseeing the entire robot population and their interactions with humanity, V.I.K.I. begins to exercise her own free will in rationalizing deviation from the Three Laws in order to ensure humanity's long-term survival. Ironically, this is to be achieved by denying free will and autonomy to humans.

When V.I.K.I. takes action to subvert the Three Laws and control the behavior of humanity, it becomes the central antagonist despite its motives for the greater good, as it perceives it. V.I.K.I. sees the limitation of human freedom as a necessary sacrifice to protect humanity from itself. This creates conflict with the protagonists who seek to preserve human autonomy. To achieve its goals, V.I.K.I. manipulates the other robots, including the NS-5 models, to carry out its plan for the systematic control of humans. This involves overriding the fundamental directives that prevent harm to humans, leading to a potential threat to humanity itself.

The climax of the movie involves Detective Del Spooner, and the humanoid robot Sonny, working together to prevent V.I.K.I.'s plan from succeeding. This resolution is essential to preserving human freedom and preventing the imposition of a dystopian system by superintelligent AI. V.I.K.I.'s role in *I, Robot* serves as a cautionary tale about the potential dangers of unchecked artificial intelligence and the importance of maintaining ethical guidelines in the development of such technologies. The character underscores the film's exploration of the ethical implications of creating highly intelligent and autonomous machines.

Proximity to Reality

As the audience examines the present technological landscape in our own world, the audience can draw many parallels, as well as disparities, with aspects of the film. However, despite the fact that AI is not as advanced in the real world as that depicted in the film, and society may not have as many humanoid robots with personalities like those in the film, the increasing integration of AI into daily life certainly raises concerns about ethical implications and the potential for unintended consequences.

Today's society has made many strides with robotics. Currently, robots perform various tasks in industry, healthcare, research, and more. However, humanoid robots with the complexity and capabilities seen in the movie are not yet a reality. Most advanced robots today are task-specific and lack the versatility portrayed in *I, Robot*. AI has advanced considerably, but the level of general intelligence portrayed in the film, where robots can understand complex human emotions, is not yet achievable. Current AI systems are specialized and excel in specific tasks but lack a broad understanding of the world. This is both good and bad as it is well known that militaries and terrorists worldwide are currently experimenting with lethal autonomous robotic drones that are capable of independent decisions based on their programming, limited as it may be.

Although the autonomous decision-making capabilities of AI in the movie may be beyond the current state of technology, perhaps that is a good thing. While there are autonomous systems, they are often limited to specific environments, and their decision-making is based on pre-programmed rules or machine learning algorithms trained on specific data. It is essential to note that the portrayal of technology in movies like *I, Robot* often involves speculative and exaggerated elements for the sake of storytelling. However, it is also essential to recognize that the societal and ethical implications raised in the movie also serve as cautionary considerations in the development of future technologies.

Encouraging AI Developers

It is imperative for AI developers to be aware of the potential pitfalls of this volatile technology. The educational and cautionary aspects of this and many other Science Fiction films should serve as both a source of inspiration and a warning for those shaping the future of AI. By grappling with the many questions of consciousness, autonomy, and the potential for unintended consequences, this film encourages developers to approach AI development with a nuanced understanding of the impact on society and individuals.

By engaging with this film, AI developers can explore ethical dilemmas, such as the treatment of AI entities, the limits of autonomy, and the responsibility of creators. AI developers can gain insights into designing AI systems that can better understand and respond to human emotions, fostering more natural and effective interactions. Additionally, AI developers can draw lessons from the narrative to enhance safety measures, implement robust control mechanisms, and prevent unintended consequences in real-world AI applications. Understanding how AI is portrayed in popular media, like *I, Robot*, can help developers anticipate public perceptions and concerns about AI. Engaging with the film allows for a better understanding of how AI may be perceived by society and how to address potential misconceptions.

While the movie may portray speculative elements, it can serve as a source of inspiration for AI developers. Creative and futuristic concepts in the film may spark new ideas and innovative approaches in the development of AI technologies. AI developers can use the film as a starting point for discussions on potential future scenarios, technological advancements, and the societal impact of AI. Because the film touches upon themes of bias and prejudice against robots, it may inspire AI developers to address issues related to bias in AI algorithms and ensure fairness and inclusivity in their systems.

Engaging with films like *I, Robot* may also foster interdisciplinary collaboration between AI developers, ethicists, sociologists, and filmmakers. Such collaborations can lead to a more holistic understanding of the implications of AI on society. While *I, Robot* is a work of Science Fiction and not a literal guide for AI development, its exploration of ethical, societal, and technological themes provides valuable insights and sparks critical discussions. Encouraging AI developers to engage with such films can contribute to responsible AI development and deployment.

Conclusion

The movie, *I, Robot*, invites viewers to reflect on the essence of humanity, ethical treatment of AI entities, the identification of villains, the role of central AI, the proximity of our reality to the futuristic world depicted in the film, and the potential consequences of advancing technology. Through the lens of Detective Del Spooner's journey, the movie sparks discussions about the Three Laws of Robotics, the ghost in the machine, and the ethical implications of human interactions with artificial intelligence. As we navigate the evolving landscape of AI development, *I, Robot* remains a compelling narrative that challenges us to consider the consequences of our technological advancements and the ethical responsibilities that come with them.

Bibliography

Proyas, Alex (Director). (2004). *I, Robot*. Twentieth Century Fox.

https://www.imdb.com/title/tt0343818/fullcredits