**‘I, Robot’, an essay discussing the portrayal of Artificial Intelligence (A.I.) in media when compared to recent A.I. Developments in Utility-Based Agents:**

Intro:

Artificial Intelligence (A.I.) refers to the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition and decision-making.

The idea of A.I. has been explored as a concept for a very long time, however it was not ‘officially’ coined until a conference in 1956.

As media such as film developed in complexity and capability, in part thanks to advances in computer and graphical technologies, so too did the portrayal of A.I. These portrayals include the omnicidal machine gestalt ‘Skynet’ from the ‘Terminator’ franchise, the curious and introspective android ‘Lt. Data’ from ‘Star Trek: Generations’, the iconic shipboard intelligence ‘Hal 9000’ from ‘2001: A Space Odyssey’, and the servile ‘Nestor-Series (NS-X)’ of Robotic servants found in this essay’s focus; ‘I, Robot’.

‘I, Robot’ is a science-fiction film released in 2004, depicting a near-future dystopian version of Chicago in 2035, where humanoid robots serve humanity, which is purportedly protected by “the Three Laws of Robotics”, a schema hard-coded into all robots intended as a safety feature.

‘I, Robot’ has been chosen as it provides a good example of a Utility-Based agent, subordinate to a more complex A.I. system that is beyond the scope of this essay.

Main:

A Utility-Based Agent is a kind of rational agent, which can make decisions based on environmental perceptions through its sensors, and take actions through the use of actuators in response to, or in anticipation of changes to its perceived environment. In a Utility-Based Agent, the information about the environment gathered by its sensors is resolved into the current world-state and agent-state. The world-state may be updated by the Agents own understanding of how the world evolves, i.e. an object in motion will remain in motion unless acted upon. The Agents current state, along with the world state can be further changed by the Agents available actions. Utility, as defined by a utility function, is a measure of action induced state change in either or both of the Agent and the world. Each action available to the Agent therefore has a given utility to the Agent. The Agents state is a measure of how “happy” it is, or how close to a given goal it is. A Utility-Based Agent thus evaluates the impact of its potential action(s) and plots a path to maximise utility and thus Agent happiness, in pursuit of a goal.

A Finite State Machine is a mathematical model that represents a system that can be in one of a finite number of states at any given time. It consists of a set of states, a set of input events or actions, a set of output actions, and a set of rules that determine how the machine moves from one state to another based o the input events it receives.

In the science-fiction film ‘I, Robot’, the ‘Nestor-Series’ of robots, the NS-4’s and NS-5’s, are examples of advanced A.I. systems. The robots can perform various tasks, ranging from playing chess to driving a car, without any human intervention. While also observing the overarching “Three-Laws of Robotics”.

The “Three Laws of Robotics” are defined in the movie as:

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.

The “Three Laws of Robotics” are purportedly ‘hard-coded’ into all the robots at the lowest level and as such, can be considered as a Finite State Machine, that is constantly evaluating whether any humans in its vicinity are at risk of coming to “harm”, before moving through to evaluating each subsequent law. This is seen in the movie when many NS-4’s are seen moving to protect the movies protagonist, exclaiming “Human in Danger!”, and are then disassembled by hostile NS-5’s, thus satisfying the 1st law, as through their action they prevented the human from coming to harm.

Each NS robot has a positronic brain which forms the core of the A.I. system and can be considered to act as multiple inter-dependant Agents as well as, in some regards, a Finite State Machine. Notably regarding ‘the Three Laws of Robotics’. For the purposes of this essay, the NS robot’s positronic brain is abstracted as the source of the robot’s utility function, the sub-agent(s) action (or movement) ‘library’ and overall agent goals. While the robot’s body is considered as the actual Utility-Based Agent to be discussed in this essay.

While technical specifications of the NS robots aren’t explored in any real depth, it can be inferred that the robot’s bodies have sensors providing; sense of sight including a large part of the electro-magnetic spectrum (to see, identify and asses any humans present, then it’s environment), sense of hearing (to hear human commands, or humans in distress), sense of touch (so they don’t apply excessive force when handling a human), and likely concealed additional sensors to provide some level of full situational awareness around the robot (low-power radar, sonar, etc).

As for actuators, the NS robots have two arms, with hands, capable of hydraulic press like grip strength, two legs, capable of launching the robot tens of feet into the air, and lastly a head containing the positronic brain, primary optical sensors and a speaker that serves as the robot’s mouth. The arms are mounted to the robot’s torso, while the legs are mounted to the robot’s pelvis, the NS-4’s have an intermediary abdominal component that links the torso and pelvis, while the NS-5’s have a more human structural skeleton.

‘I, Robot’ was released in 2004, and depicted a near-future idea of what robots and by extension A.I. agents, might become. At the time of the movie’s release the robots as shown were reasonably realistic; performing feats that, while impressive, remained within the realms of physical possibility, even while being completely unfeasible to recreate on many levels. For example, the autonomy displayed by the NS robots was and still is far beyond the ability of modern A.I. systems, additionally the robot’s power-source is never mentioned, and the ‘positronic brain’ that drives the robots remains a purely fictional technology.

The depiction is considered as “reasonably” realistic since the robots are mostly seen performing human tasks, with the occasional display of robotic strength. The main point of fiction is the intelligence and autonomy displayed by the robots, which was completely beyond the realms of possibility in 2004.

When compared to the present day – March 2023 at time of writing – however, while the autonomy shown by the NS robots remains unreproducible, the physical capabilities are considerably closer to being realised.

The real-world contemporary to the NS robots at present would be Boston Dynamic’s AtlasTM; “Atlas is a research platform designed to push the limits of whole-body mobility.” -[Atlas™ | Boston Dynamics](https://www.bostondynamics.com/atlas). Atlas is a humanoid robot, capable of navigating within a managed environment and performing simple tasks correctly, some of the time. As it is still in active development and, a research platform to boot, Atlas still has trips and falls. It is possible a finite state machine is responsible for evaluating whether the robot has lost control of its situation and thus moves to cut hydraulic power in an attempt to minimise damage to the robot, although this is unclear.

The Atlas platform lacks the autonomy of the NS robots and requires large amounts of human intervention to achieve its impressive displays of mobility. It is not infeasible however, to consider a more complex A.I. system providing goals, such as move from point A to point C, via route B, to the Atlas Robot and, having its onboard A.I. system, viewed primarily as a Utility-Based Agent, chose a sequence of movements, such as walking and jumping, to move from point A to point C, via route B.

Conclusions:

While there have been significant advances in the field of A.I., we are still far from creating autonomous robots that can perform complex tasks without human supervision. Current A.I. systems often require substantial amounts of data and human input to function accurately.

Other recent developments in A.I. have been focused on developing machines with specific functional applications. Utility-Based agents are designed to carry out tasks based on pre-defined datasets and algorithms. This approach to A.I. systems, in concert with other technologies has been applied towards making various tasks more efficient and completely automating simpler tasks to improve productivity.

Overall, while the state of A.I. in ‘I, Robot’ represents a vision of what A.I. technology could look like in the future, it is not an accurate reflection of current developments in the field, which has had a narrower focus when compared to the advanced and largely autonomous systems depicted in the film.

Speculatively, looking at current developments and trends in A.I. systems and robotics, it is possible that within the next decade, robots with similar capabilities to those seen in ‘I, Robot’, such as improved versions of the Atlas platform, could be remotely directed by large supercomputers or computational clusters running multiple interdependent A.I. systems to achieve a similar result as seen in the movie. This prediction is based on recent strides in natural language A.I. models, such as ChatGPT; The aforementioned Atlas Platform; Existing pathfinding/autonomous driving (autopilot) technologies such as those seen in Tesla model cars; And recent improvements in text-to-speech and voice replication techniques, where it is becoming increasingly easier to replicate the voices of specific people, such as those of US presidents, as seen in “US Presidents play <Insert Game>” videos on YouTube. For example: [US Presidents Play New Super Mario Bros. Wii - YouTube](https://www.youtube.com/watch?v=GZKgYsHGm-w)

References:

Boston Dynamics (2019). *Atlas | Boston Dynamics*. [online] Bostondynamics.com. Available at: <https://www.bostondynamics.com/atlas>.

Wikipedia Contributors (2019). *Three Laws of Robotics*. [online] Wikipedia. Available at: <https://en.wikipedia.org/wiki/Three_Laws_of_Robotics>.

www.youtube.com. (n.d.). *US Presidents Play New Super Mario Bros. Wii*. [online] Available at: https://www.youtube.com/watch?v=GZKgYsHGm-w [Accessed 31 Mar. 2023].