

Robots and AI

by Roger Plant, FSPE

AI, Robots and Humans: An Ethical Dilemma.

Recent developments in AI have led me to think up the following dilemma. If an AI device passes the Turing test and is therefore indistinguishable from a human, does that mean it has the same rights as a human? Does it have the right not to be switched off, terminated or “killed off” as a human would? I coin the phrase Human Equivalent AI (HEAI) to describe such a device.

Developed by Alan Turing in 1950, the Turing test is “a test of a machine’s ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. Turing proposed that a human evaluator would judge natural language conversations between a human and a machine designed to generate human-like responses.”¹

Also during the 1950s, Isaac Asimov developed the Laws of Robotics designed to protect humans from rogue robots. The Three Laws of Robotics (often shortened to The Three Laws or known as Asimov’s Laws).²

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.

Later amended by adding the “zeroeth” law:

0. A robot may not harm humanity, or, by inaction, allow humanity to come to harm.

But can an HEAI claim to be human and therefore not be subject to Asimov’s Laws?

In the film *Ex_Machina* we see a humanoid robot passing a Turing test. That Human Equivalent AI (HEAI), in the film called *AVA*, after passing the Turing test, demands to have the same rights as a human. But its creator refuses to let it out of secure confinement. In the same way as a human would, *AVA* seriously objects to being so confined and turns against its creator in order to escape its confinement.

So the question remains, “Does an HEAI have the same rights as a human?” I consider this question as important as Turing’s original test and Asimov’s Laws of Robotics. Turing developed his test in order to classify an AI device as indistinguishable from a human and therefore logically “not a robotic device.”

As an afterthought to this question, it would be easy to switch off an HEAI device if it were a grey box plugged into a wall socket, but an HEAI built to physically resemble a human, and therefore able to defend itself, may be much harder to disable.

Also in the future, any HEAI may well be able to claim the same legal rights as a human and therefore have the legal right to sustenance to keep it “alive.” Humans are not allowed to starve another human to death, and an HEAI may be able to claim the legal right not to be starved of the energy source that keeps it going.

The main critical point is that an HEAI is *intellectually* indistinguishable from a human.

Its physical form or means of reproduction or manufacture are, in my opinion, not really important. Biological reproduction is common to all life forms, but the critical point that separates humans from all other life forms, even the highest levels of animals, is intellectual ability. It is that intellectual ability that allows the HEAI to pass the Turing test

and make it indistinguishable from a human. That is the critical point that makes it an HEAI and raises the question of rights.

One simple answer I can see is that if Azimov's Laws of Robotics cannot be applied to an HEAI, then human laws can. In other words, an HEAI would be subject to the full range of human laws which should overcome the ethical dilemma.

To highlight the classification I apply to HEAI, we have to consider various robotic devices. There are humanoid robots, or androids as they are commonly called, that have human-like structure such as Atlas and similar devices built by Boston Dynamics

(<https://www.bostondynamics.com/atlas>), and of course the myriad of human-shaped devices that are produced currently as advanced toys. But I would not expect such devices to pass a Turing test, at least not today! An HEAI could be either humanoid in appearance or simply a box-like device, either of which may be able to pass a Turing test.

It is the devices that can pass the Turing test, and therefore be intellectually equivalent to humans, that I call HEAI and whose (its? :-)) rights we should be considering.

To follow this discussion in more detail, please see http://the1000.ning.com/group/ai?xg_source=activity.

NOTES

1. Wikipedia contributors, "Turing test," Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/w/index.php?title=Turing_test&oldid=822305924.
2. Wikipedia contributors, "Three Laws of Robotics," Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/w/index.php?title=Three_Laws_of_Robotics&oldid=819461502.

■ ■ ■
“Your theory is crazy,
but it's not crazy enough to be true.”
— Niels Bohr

■ ■ ■
“Man will occasionally stumble over the truth,
but most of the time he will pick himself up
and continue on.”
— Winston Churchill