

Technology

Can Robots Make Ethical Decisions?

IN COOPERATION WITH:



By Bill Christensen, Technovelgy.com
posted: 16 September 2009 07:50 pm ET

Robots and computers are often designed to act autonomously, that is, without human intervention. Is it possible for an autonomous machine to make moral judgments that are in line with human judgment?

This question has given rise to the issue of machine ethics and morality. As a practical matter, can a robot or computer be programmed to act in an ethical manner? Can a machine be designed to act morally?

Isaac Asimov's famous fundamental Rules of Robotics are intended to impose ethical conduct on autonomous machines. Issues about ethical behavior are found in films like the 1982 movie *Blade Runner*. When the replicant Roy Batty is given the choice to let his enemy, the human detective Rick Deckard, die, Batty instead chooses to save him.

A recent paper published in the International Journal of Reasoning-based Intelligent Systems describes a method for computers to prospectively look ahead at the consequences of hypothetical moral judgments.

The paper, *Modelling Morality with Prospective Logic*, was written by Luís Moniz Pereira of the Universidade Nova de Lisboa, in Portugal and Ari Saptawijaya of the Universitas Indonesia. The authors declare that morality is no longer the exclusive realm of human philosophers.

Pereira and Saptawijaya believe that they have been successful both in modeling the moral dilemmas inherent in a specific problem called "the trolley problem" and in creating a computer system that delivers moral judgments that conform to human results.

The trolley problem sets forth a typical moral dilemma; is it permissible to harm one or more individuals in order to save others? There are a number of different versions; let's look at just these two.

Circumstances There is a trolley and its conductor has fainted. The trolley is headed toward five people walking on the track. The banks of the track are so steep that they will not be able to get off the track in time.

Bystander version Hank is standing next to a switch, which he can throw, that will turn the trolley onto a parallel side track, thereby preventing it from killing the five people. However, there is a man standing on the side track with his back turned. Hank can throw the switch, killing him; or he can refrain from doing this, letting the five die.

Is it morally permissible for Hank to throw the switch?

What do you think? A variety of studies have been performed in different cultures, asking the same question. Across cultures, most people agree that it is morally permissible to throw the switch and save the larger number of people.

Here's another version, with the same initial circumstances:

Footbridge version Ian is on the footbridge over the trolley track. He is next to a heavy object, which he can shove onto the track in the path of the trolley to stop it, thereby preventing it from killing the five people. The heavy object is a man, standing next to Ian with his back turned. Ian can shove the man onto the track, resulting in death; or he can refrain from doing this, letting the five die.

Is it morally permissible for Ian to shove the man?

What do you think? Again, studies across cultures have been performed, and the consistent answer is reached that this is not morally permissible.

So, here we have two cases in which people make differing moral judgments. Is it possible for autonomous computer systems or robots to come to make the same moral judgments as people?

The authors of the paper claim that they have been successful in modeling these difficult moral problems in computer logic. They accomplished this feat by resolving the hidden rules that people use in making moral judgments and then modeling them for the computer using prospective logic programs.

Ethical dilemmas for robots are as old as the idea of robots in fiction. Ethical behavior (in this case, self-sacrifice) is found at the end of the 1921 play *Rossum's Universal Robots*, by Czech playwright Karel Capek. This play introduced the term "robot".

Science fiction writers have been preparing the way for the rest of us; autonomous systems are no longer just the stuff of science fiction. For example, robotic systems like the Predator drones on the battlefield are being given increased levels of autonomy. Should they be allowed to make decisions on when to fire their weapons systems?

The aerospace industry is designing advanced aircraft that can achieve high speeds and fly entirely on autopilot. Can a plane make life or death decisions better than a human pilot?

The H-II transfer vehicle, a fully-automated space freighter, was launched just last week by the Japan's space agency JAXA. Should human beings on the space station rely on automated mechanisms for vital needs like food, water and other supplies?

Ultimately, we will all need to reconcile the convenience of robotic systems with the acceptance of responsibility for their actions. We should have taken all of the time that science fiction writers have given us to think about the moral and ethical problems of autonomous robots and computers; we don't have a lot more time to make up our minds.

- 5 Reasons to Fear Robots
- Implanted. Enhanced. Invaded?: Human-Robot Mergers
- More Robot News

This Science Fiction in the News story used with permission of Technovelgy.com.

[Space](#) | [Animals](#) | [Health](#) | [Environment](#) | [Technology](#) | [Culture](#) | [History](#) | [Strange News](#) | [Videos](#) | [Galleries](#) | [Hot Topics](#) | [Community](#) | [Tech Reviews](#) | [Store](#)

 **IMAGINOVA.**

 LIVE SCIENCE

 SPACE

 NEWSrama

 SPACE NEWS

 Starry Night

 ORION

[Site Map](#) | [About Us](#) | [Contact Us](#) | [Privacy Policy](#) | [Terms and Conditions](#) | [Advertise with Us](#) | [DMCA/Copyright](#)

© Imaginova Corp. All rights reserved.