



Online Ethics Center
FOR ENGINEERING AND SCIENCE

Artificial Intelligence & Robotics Subject Aid

Author(s)

Rachelle Hollander

Year

2016

Description

A short guide to some key resources and readings on the topic of AI and Robotics.

Body

Robotics is a branch of science and engineering that develops particular kinds of automated machines for a very wide variety of purposes. The fields of mechanical engineering, electrical engineering, and computer science are most directly involved with the design, construction, operation, and application of robots as well as computer systems for their control, sensory feedback, and information processing. Robots often contain sophisticated computer systems to guide their interactions with the environment. Some robots exhibit abilities that resemble those of human beings or other animals – mobility, control and sensory feedback. They can be similar in terms of information processing and decision making, including learning and problem solving. Machines are said to have artificial intelligence (AI) when they exhibit these attributes – AI machines are able to sense or assess their environment to achieve some goal. These machines are typically designed to perform domestic, commercial, or military tasks. They are often built to do jobs that are hazardous,

repetitive, or boring to people. With improvements in AI, an increasing number of robots are programmed to behave autonomously, without human intervention or control over their actions.

In computer science, an ideal "intelligent" machine is a flexible rational agent that perceives its environment and takes actions that maximize its chance of success at some goal. Some philosophers would contest the use of the term "rational" because it implies consciousness. Without resolving that dispute, scientists and engineers working in robotics apply the term "artificial intelligence" when a machine mimics "cognitive" functions that are normally associated with a human mind. AI in computer science applies to machines that mimic human cognitive functions such as learning and problem solving. Current AI capabilities include understanding human speech, competing at a high level in such strategic game systems such as Chess and Go, self-driving cars, and interpreting complex data.

Development and use of robots can result in both benefits and harms to human beings, whether they are or are not equipped with AI. Ethical, legal, and social issues related to robots will continue to emerge in realms such as health care, manufacturing, and warfare. The potential for tremendous change in social and economic arrangements as developments in AI and robotics evolve and the difficulties that would pose to norms of behavior and in society makes this a topic of great interest among scientists, engineers, humanities and legal scholars, policymakers, and the public.

This entry was drawn from Wikipedia definitions for Robotics, Artificial Intelligence, and Robot Ethics:

Robotics. Wikipedia; accessed Nov. 1, 2016.

<https://en.wikipedia.org/wiki/Robotics>

Artificial Intelligence. Wikipedia; accessed Nov. 1, 2016.

https://en.wikipedia.org/wiki/Artificial_intelligence

Robot Ethics, Wikipedia; accessed Nov. 2, 2016,

https://en.wikipedia.org/wiki/Robot_Ethics

Subject Overviews

Lin, Patrick, Keith Abney, and George A. Bekey, Editors. 2011. *Robot ethics: the ethical and social implications of robotics*. Cambridge, MA: MIT Press.

Starting with an overview of the issues and relevant ethical theories, the topics move to the possibility of programming robot ethics, the ethical use of military robots in war, to legal and policy questions, including liability and privacy concerns. The contributors then turn to human-robot emotional relationships, examining the ethical implications of robots as sexual partners, caregivers, and servants. Finally, they explore the possibility that robots, whether biological-computational hybrids or pure machines, should be given rights or moral consideration.

Wallach, Wendell and Colin Allen. 2009. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press. Extended abstract of the book by Tony Beavers at https://philosophynow.org/issues/71/Moral_Machines_Teaching_Robots_Right_from_Wrong. Accessed July 25, 2016.

Computers are already approving financial transactions, controlling electrical supplies, and driving trains. Soon, service robots will be taking care of the elderly in their homes, and military robots will have their own targeting and firing protocols. Colin Allen and Wendell Wallach argue that as robots take on more and more responsibility, they must be programmed with moral decision-making abilities, for our own safety. Engineers exploring design strategies for systems sensitive to moral considerations in their choices and actions will need to determine what role ethical theory should play in defining control architectures for such systems. See chapter two for a framework for moral agency and chapter three for a discussion of the specific concerns for human freedoms and responsibility that such autonomous, morally sensitive agents might raise.

Sullins, John. Summer 2016 Edition. "Information Technology and Moral Values", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/sum2016/entries/it-moral-values/>. First published June 2012.

Section 2.4 discusses Future Concerns, under the headings “Acceleration of Change,” “Artificial Intelligence and Artificial Life,” and “Robotics and Moral Values.”

Noorman, Merel. Summer 2016 Edition. "Computing and Moral Responsibility", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). See <http://plato.stanford.edu/archives/sum2016/entries/computing-responsibility/>. First published June 2012.

In contemporary life, technologies mediate much human interaction, including of course those with moral content. This raises questions and complications about the ascription of moral responsibility. See particularly section 2, “Can computers be moral agents?” and 3, “Rethinking the concept of moral responsibility.”

Arkin, Ronald C. 2009. Ethical robots in warfare. *IEEE Technology and Society Magazine*. 28(1): 30-33. DOI: 10.1109/MTS.2009.931858

This article argues that robots not only can be better than soldiers in conducting warfare in certain circumstances, but they also can be more humane in the battlefield than humans. As robots can be built that do not exhibit fear, anger, frustration, or revenge, and that ultimately (and the key word here is ultimately) behave in a more humane manner than even human beings in these harsh circumstances and severe duress. People have not evolved to function in these conditions, but robots can be engineered to function well in them.

Casner, Stephen M., Edwin.L. Hutchins, and Don Norman. 2016. “The challenges of partially automated driving.” *Communications of the ACM*, 59(5), 70-77. <http://cacm.acm.org/magazines/2016/5/201592-the-challenges-of-partially-automated-driving/fulltext>

While many assume that the march to the automated car is underway, adequate attention is not being paid to the problems that are arising and will arise in the transition. These problems are significant, raising life and death issues. The authors point out four levels of automation in which it seems that only the final level would preclude any role for drivers. In the interim, the transitional problems should be addressed.

Lin, Patrick. 2013. The ethics of autonomous cars." *The Atlantic*. October 8. Online archive at <http://www.theatlantic.com/technology/archive/2013/10/the-ethics-of-autonomous-cars/280360/>

The author points to many circumstances in which human drivers exercise judgment in ways that may be difficult to impossible to duplicate in code for self-driving vehicles and where ethics, law, and policy may not align. He identifies numerous conditions which may have specific implications for how autonomous vehicles should behave which have not received consideration arguing that the ethical issues should be addressed in an anticipatory fashion.

Sparrow, Robert, and Sparrow, Linda. 2006. In the hands of machines? The future of aged care. *Minds and Machines* 16: 141-161.

This paper surveys and assesses the claims made on behalf of robots in relation to their capacity to meet the needs of older persons paying particular attention to the social and ethical implications of the introduction of robots. It argues that their use is likely to decrease the amount of human contact and that this result is unethical. It places its view in the context of broader social attitudes towards older persons and proposes the need for a deliberative process involving them as a test for the ethics of using robots in aged care.

Policy or Guidance

Chameau, Jean-Lou, William F. Ballhaus, and Herbert S. Lin, Editors. National Research Council. 2014. Chapter 3, Section 3.1. "Robotics and Autonomous Systems." in *Emerging and Readily Available Technologies and National Security - A Framework for Addressing Ethical, Legal, and Societal Issues*. 79-92. Washington, DC: National Academies Press. <http://www.nap.edu/read/18512/chapter/5>. Accessed July 24, 2016.

The discussion focuses on the ethical, legal, and social implications in the development and use of these systems for military purposes.

Human Rights Watch. 2012. Losing Humanity, The Case against Killer Robots. Posted November 19, 2012. Accessed November 1, 2016.
<https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>.

The primary concern of Human Rights Watch and IHRC is the impact fully autonomous weapons would have on the protection of civilians during times of war. This report analyzes whether the technology would comply with international humanitarian law and preserve other checks on the killing of civilians. It finds that fully autonomous weapons would not only be unable to meet legal standards but would also undermine essential non-legal safeguards for civilians. The research and analysis strongly conclude that fully autonomous weapons should be banned and that governments should urgently pursue that end.

National Research Council. *Interfaces for Ground and Air Military Robots: Workshop Summary*, 2005. <http://www.nap.edu/catalog/11251/interfaces-for-ground-and-air-military-robots-workshop-summary>. Pages 31-32

In the final section on key issues this workshop summary identifies ethical and social concerns – one is the likelihood that those who first encounter the new technology may not be those expected; they may be children or, on the other hand, they may be persons whom one does not want to have access. Other issues identified are problems that come with autonomous function, the potential diminution of moral responsibility, and privacy breaches.

Bibliography

Online Ethics Center. 2020. “Robot Morality and Artificial Agency Bibliography”. <https://onlineethics.org/cases/robot-morality-and-artificial-moral-agency-bibliography> Accessed July 25, 2016.

This extensive, annotated bibliography contains book and journal article citations as well as web resources on the ethics of artificial intelligence and machine intelligence.

Notes

Reviewed by Jason Borenstein, November 2, 2016

Rights

Use of Materials on the OEC

Resource Type

Bibliography

Parent Collection

OEC Subject Aids

Topics

Artificial Intelligence and Robotics

Controversies

Discipline(s)

Computer, Math, and Physical Sciences

Social and Behavioral Sciences

Computer Engineering

Electrical Engineering

Authoring Institution

Online Ethics Center