

PHIL 20632/STV 20233
Robot Ethics
TTh 11:00-12:15
203 DeBartolo

Spring 2017

Prof. Don Howard
308 Malloy Hall
Tel: 631-7547
dhoward1@nd.edu
Office Hours: TTh 2:00-3:00

Texts: There are two required books for this course:

Ronald C. Arkin. *Governing Lethal Behavior in Autonomous Robots*. CRC Press, 2009.

Patrick Lin, Keith Abney, and George A. Bekey, eds. *Robot Ethics: The Ethical and Social Implications of Robotics*. MIT Press, 2012.

A few additional readings will be distributed directly to the class in electronic form.

Course Web Site: www.nd.edu/~dhoward1/ - Then follow the obvious link.

Requirements: There will be three components in the computation of your final grade for the course:

(1) **Discussion Papers** (60 %). Each student will be required to submit three discussion papers, each a minimum of five pages, on topics to be worked out in consultation with the instructor. Each of the four discussion papers will be worth 20 % of the final course grade. Papers will be graded on the basis of both content and mechanics, the latter counting for approximately 20% of the grade on each paper.

(2) **Journals** (25 %). You will be required to keep a journal for recording your critical reactions to the assigned readings. At a minimum, you should write a one-page (≥ 300 words) critical response to each reading assignment, though you are encouraged to write more, taking advantage of this opportunity to record your thoughts on any topic related to the course and the readings. Journal entries will be graded on a random, unannounced basis several times over the course of the semester. In addition, the entire journal will be checked regularly to be sure that all required entries are included. You will submit your journal entries through the Sakai page for this course, creating a “blog” entry for each of the reading assignments. (Be sure to remember to check the option “**Only site administrators and I can see this entry**” and to “publish” your entry.) You will be graded rigorously on the extent to which you use your journal, the minimum of one page (300 words) per reading being strictly enforced. But it will be the quality of thinking manifest in your journal that will chiefly determine your grade. For more on how to keep a good journal, see the extra handout on that topic.

(3) **Class Participation** (15 %). The remaining fifteen percent of your final grade will be determined on the basis of the quality and extent of your enthusiastic and constructive participation in class discussion.

One-minute Papers: Every class session will end a few minutes early to permit you to write a so-called “one-minute paper,” in which you will write no more than two- or three-sentence answers to two questions: (a) What was the most important point covered in today’s class? (b) What issue or question was left most unclear in your mind at the end of today’s class? These one-minute papers will be required of every student at the end of every class session and will be collected at the end of class, but they will not be graded.

Attendance: No more than two unexcused absences will be permitted during the semester; for every additional unexcused absence, the student’s final course grade will be reduced by one step on a plus/minus basis. Thus, for example, for a student with three unexcused absences, a final grade of B+ becomes a B, and for a student with four unexcused absences, a final grade of B becomes a C+. A student more than fifteen minutes late for class is assumed to be absent. If you must miss a class for any reason, be sure to let me know beforehand, or as soon after the fact as possible, so as to enable me to determine whether or not to excuse the absence. I promise to be generous in allowing excused absences for legitimate purposes.

Schedule:

Date:	Topic:	Readings:
17 Jan.	Introduction - What is Robot Ethics?	
19 Jan.	Should We Do It?	Arkin, "The Case for Ethical Autonomy in Unmanned Systems." Human Rights Watch, "Losing Humanity: The Case against Killer Robots."
24 Jan.	The Push for Autonomy in Weapons Systems	Arkin, "Introduction" and "Trends toward Lethality," Chs. 1 & 2 of <i>GLBAR</i> .
26 Jan.	Autonomy in Weapons Systems: Why and Whether	Arkin, "Human Failings in the Battlefield," "Related Philosophical Thought," and "What People Think: Opinions on Lethal Autonomous Systems," Chs. 3-5 of <i>GLBAR</i> .
31 Jan.	Building Lethal Autonomous Robots, I	Arkin, "Formalization for Ethical Control," and "Specific Issues for Lethality: What to Represent," Chs. 6 & 7 of <i>GLBAR</i> .
2 Feb.	Building Lethal Autonomous Robots, II	Arkin, "Representational Choices: How to Represent Ethics in a Lethal Robot" and "Architectural Consideration for Governing Lethality," Chs. 8 & 9 of <i>GLBAR</i> .
7 Feb.	Building Lethal Autonomous Robots, III	Arkin, "Design Options" and "Example Scenarios for the Ethical Use of Force," Chs. 10 & 11 of <i>GLBAR</i> .
9 Feb.	<i>No class.</i>	
14 Feb.	Building Lethal Autonomous Robots, IV	Arkin, "A Prototype Implementation," Ch. 12 of <i>GLBAR</i> .
16 Feb.	In Defense of Autonomous Weapons	Don Howard, "In Defense of (Virtuous) Autonomous Weapons."
21 Feb.	Technology and Ethics: The Lay of the Land	Patrick Lin, "Introduction to Robot Ethics" and George Bekey, "Current Trends in Robotics: Technology and Ethics"

23 Feb.	Technology and Ethics: Frameworks for Ethical Decision Making	Keith Abney, "Robotics, Ethical Theory, and Metaethics: A Guide for the Perplexed"
27 Feb.	<i>First Discussion Paper Due</i>	
28 Feb.	Design and Programming	Colin Allen and Wendell Wallach, "Moral Machines: Contradiction in Terms or Abdication of Human Responsibility" and James Hughes, "Compassionate AI and Selfless Robots: A Buddhist Approach"
2 Mar.	Ethics Programming for Artificial Autonomous Moral Agents	Don Howard and Ioan Muntean, "Artificial Moral Cognition: From Functionalism to Autonomous Moral Agents."
7 Mar.	Law and Robotics	Richard O'Meara, "Contemporary Governance Architecture Regarding Robotic Technologies: An Assessment" and Peter M. Asaro, "A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics"
9 Mar.	Law and Robotics	M. Ryan Calo, "Robots and Privacy"
13-17 Mar.	<i>Spring Break</i>	
21 Mar.	Love, Sex, and Robots	Matthias Scheutz, "The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots" and David Levy, "The Ethics of Robot Prostitutes"
23 Mar.	Love, Sex, and Robots	Blay Whitby, "Do You Want a Robot Lover? The Ethics of Caring Technologies"
27 Mar.	<i>Second Discussion Paper Due</i>	
28 Mar.	Medicine and Care	Jason Borenstein and Yvette Pearson, "Robot Caregivers: Ethical Issues across the Human Lifespan" and Noel Sharkey and Amanda Sharkey, "The Rights and Wrongs of Robot Care"
30 Mar.	Medicine and Care	Steve Petersen, "Designing People to Serve."
4 Apr.	"Star Trek: The Next Generation," episode 2.9, "The Measure of a Man"	Video in class
6 Apr.	Rights and Ethics	Rob Sparrow, "Can Machines Be People? Reflections on the Turing Triage Test" and Kevin Warwick, "Robots with Biological

		Brains”
11 Apr.	Rights and Ethics	Anthony Beavers, “Moral Machines and the Threat of Ethical Nihilism”
13 Apr.	Robotics and Prosthetics	Class visit by Eric Earley and Max Shepherd from the Center for Bionic Medicine
18 Apr.	Codes of Ethics	Riek and Howard, "A Code of Ethics for the Human-Robot Interaction Profession"
20 Apr.	Beyond the Advancing Edge of Technology: I. Self-Driving Cars	Patrick Lin, “The Ethics of Saving Lives With Autonomous Cars Are Far Murkier than You Think”
25 Apr.	Beyond the Advancing Edge of Technology: I. Self-Driving Cars	Mark Mills and Don Howard, “Punch the Accelerator on Self-Driving Cars.”
27 Apr.	Beyond the Advancing Edge of Technology: II. The AI Singularity	Ray Kurzweil, “The Coming Singularity.” Neil deGrasse Tyson and Ray Kurzweil, “2029: Singularity Year.”
2 May	Beyond the Advancing Edge of Technology: II. The AI Singularity	Paul Allen, “The Singularity Isn’t Near”; Ray Kurzweil, “Don’t Underestimate the Singularity”
5 May	<i>Third Discussion Paper Due</i>	