Humans have always used the resources and tools around them to adapt their environment to themselves and enhance or improve their lives: from glasses to Google glass, brooms to roombas, and even increasingly to robots. Throughout the Twentieth century, popular media from theater, to books, and films have explored a range of potential consequences for our lives with futuristic technology including robots. For the most part, representations of robots in books and media are popular instances of express collective fears about machines taking over the realm of the human, our tools getting out of our control, and the potential benefits and dangers of the things we make. Today, robots of various kinds are widely used in factories and increasingly in our homes; most scientists in the field of robotics predict we will all have them in our houses within 20 years. Meanwhile scholars from a range of disciplines are thinking through the ethical, social, and even theological implications of this new technology, always with the doomsday scenario of fiction lingering in the background. Through engagement with science fiction in literature and film as well as emerging science non-fiction in contemporary robotics laboratories, this set of readings is designed to help you think together about the relationship between humans and humanoid machines.

I designed this as a Freshman seven week research seminar which I know almost all of you are familiar with, at least in the deepest recesses of your memory. In these seminars students are able to stretch their wings a bit, examining a topic that interests them in greater depth using the close reading, critical thinking, and academic writings skills they have been honing all year to write a research paper. My aims in crafting this seminar were twofold: 1) to help freshmen put the traditional texts we read in dialogue with issues that are emerging in their lifetimes and 2) to put literature, philosophy, and science into dialogue. This topic does both naturally. As you can see through the title, the underlying premise of the seminar is that examining the implications of this new technology in human life forces us to think about what it means to be human. And as you will see, once you have robots on your radar, that they are everywhere and burgeoning.1

The five sessions outlined here each include a text and at least one film, and sometimes other media. I have provided links to anything available on the web legally. Please note: this set of readings and other media barely scratches the surface, there are many other topics that could be covered here (warfare, medicine, etc.). Enjoy!

**Texts:**


---

1 In fact, for those of you in the Chicago region the Museum of Science and Industry has an excellent robot exhibit called “Robot Revolution” running until January 3, 2016. You may want to plan a trip with your reading group! Information can be found here: [http://www.msichicago.org/whats-here/exhibits/robot-revolution/](http://www.msichicago.org/whats-here/exhibits/robot-revolution/).
Films:
Wall-e (2008)
First Discussion: Talking Robots

The term robot is used in a range of ways each emphasizing different capacities. According to the Oxford English Dictionary, robot means “A machine capable of carrying out a complex series of actions automatically, especially one programmable by a computer.” While science fiction author and Isaac Asimov, defined them as “an artificial object that resembles a human being” and later “machine + computer”. A recent article in the Wall Street Journal defined them as, “any automated process that substitutes machines for people”. While each of these definitions hold for the wide range of machines called robots today, they each emphasize a different feature or capacity, intelligence, physical form, job or capability. In class we came up with the following features to characterize robots, feel free to rethink of modify as you read and as you discuss throughout the year:

- machines/tools/technology we make
- that perform automated processes (programed, learn, think?)
- that do labor/work humans can’t do, can’t do safely, or that is monotonous
- may or may not resemble humans

The notion of a robot is both new and not so new. In Homer’s *The Iliad* there is an account of non-human, female helpers made of gold who can think and use their muscles to weave and do work. They might well be the earliest account of a robot. Thinking technologically, the 1700s was a heyday for automatons, clockwork toys and displays that could be found throughout Europe in town centers and even in churches until reformers became nervous about these representations of God and Jesus. It seems humans have long been fascinated with the notion of mechanical creatures as helpers or entertainment.

The word “robot” comes from Czech word meaning “slaves” or “forced workers.” In Czech *robota* referred to serfs in medieval Europe whose labor was the currency for rent. It came to be used to describe any kind of hard or unpleasant work. In 1920 the playwright Karel Capek wrote a play called *R.U.R.* (*Rossum’s Universal Robots*) in which robots were created to take over the human’s labor so humans could live in leisure and comfort. Capek thought about calling them *labori* (Latin for labor) but thought it sounded to scholarly so turned to *roboti* to emphasize their enslaved state of labor. *Roboti* became robot in the first English translation 1923. In the play the robots look remarkably like humans. They develop emotions and get tired of being slaves and especially military slaves so they overthrow humanity. In this initial plot we see the core of robot fear, fascination, and fantasy of robots in a nutshell.

There are a few related terms that come up from time to time in robot literature which I want to outline at the start.

- **Android**: A robot, fully mechanical being, who resembles a human in form. Think C3PO or Commander Data.
- **Cyborg**: Combining cyber (electronic network) and organism, this term is commonly used to refer to a human whose physical abilities are extended beyond normal human limitations by mechanical elements built into the body. Think Battle Star Galactica, or yourself (if you wear glasses, have a pace maker, hearing aid, or prosthetic leg or arm).

---

In contemporary social and philosophical theories of transhumanism any enhancement to the body by technology can be considered cyborg (glasses, google glasses, hearing aid, pace maker, prosthetics). Some argue that this is the next phase of human evolution to enhance ourselves with technology, sometimes referred to as post-humanism. Others argue this may be where we are going, but it is terrible, the death of humanity. Either way this is a controversial usage of the term cyborg and hotly debated today. Addressing these issues would require an entirely different reading list but I wanted to alert you to this area of inquiry.

- **Artificial Intelligence**: This refers to any computer intelligence. AI is an entire field of research in its own right. While I have tried to keep this syllabus focused on robots, AI will come up from time to time as the intelligence or computer system that runs a robot. Scientists working on robots are sometimes working on AI (the brain) and sometimes on the body/interactions but all robots require both.

**Isaac Asimov (1920-1992)**

Isaac Asimov is probably the foremost robot fiction writer. Hardly any account of robots goes by without at least passing reference to Asimov. He is one of the first to really theorize a world with robots. The set of short stories that we are reading for today are some of his earliest stories and show the evolution of his own thinking about the potential ways robots would evolve and some of the potential issues that might arise for humans working with robots. Asimov had read *R.U.R* and did not like what he called the “Frankenstein complex” that most accounts of robots seemed to have. He set out to think of a world with robots who weren’t necessarily menacing, where the problems that come up were more complicated. In this he was successful. He is probably most famous for coining the term “robotics” and “positronic brain” in his short stories. He also invented the Laws of Robotics (sometimes calle Asimov’s 3 laws) in one of his earliest stories, “Runaround” (1941), which we read for today. The 3 laws are:

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 Law.

Asimov argued that just as we humans make safety rules for tools, if we developed robots we would develop similar safety measures. These rules serves as the scaffolding for all of his stories, testing the limits of the system for robot human relations.

The readings for this session were all published separately over a period of nearly ten years, but here they hold together almost like a novel, with repeating characters and a frame of an interview with Susan Calvin the world’s foremost robopsychologist.


What Robots Teach Us About Being Human
CC Alumni Reading Group Syllabus 2015-16

Discussion questions:
- As you read through these short stories what do you notice about the ways Asimov builds on the potentiality and complications of robots? What kinds of work do the robots do? What are they like? How do they interact with humans? How do humans interact with them? What is the world Asimov builds like?
- We get glimpses of the political, legal, cultural situation from time to time, where was Asimov prescient and where was he off base (or too early)? What do you make of the human characters?
- Gender is something that stands out throughout the stories in two main ways, 1) through the persona of Susan Calvin and the reaction of those around her; 2) all the robots seem to be male. How did you experience these aspects of the short stories, most of which were written before WWII?

You may want to focus your attention on a few key stories.
- “Robbie” (1939) is Asimov’s first published robot story, published when he was 19. In the story, what is Robbie like? Asimov juxtaposes the mother and daughter’s response to Robbie and to the pet dog. How is Asimov envisioning the relationship between humans and robots?
- “Runaround,” and “Reason” are about robots gone awry, what problems do the human’s encounter? “Reason” touches on theology and creation as well as ethics? What do we learn from this story?
- In both “Catch that rabbit” and “Little lost Robot” Asimov is testing out the three laws and implications of changing them, what affect do the changes have?
- In “Liar” Herbie interacts with humans differently than the other robots we have seen? What new capacities do we see here and what are the ramifications?
- “Evidence” starts to test the issue of whether or not we can tell if someone is a robot. What new problems does this raise?
- Given that robots are banned from earth in Asimov’s world but not ours would these concerns matter today? Why or why not? Why ban robots from earth?

Thinking through the 3 laws:
- All of the examples of the things that go wrong in robots in these stories parallel human ethics, or emphasize things we assume. Asimov says in a later essay, that these can be considered the three laws of tools: 1) a tool must be safe to use; 2) a tool must perform its function, provided it does so safely; 3) a tool must remain intact, unless to do so keeps it from being safe, or performing its function (Robot Visions, “The laws of robotics”). Does this analogy work? Where might it fall short?
- Things to keep in mind as you think through the 3 laws: Does law 1 imply care for humans, mere safety, simulated care, emotion, or enslavement? Does law 2 imply responsibility on the part of humans? Does it imply rights for robots? Does law 3 imply care for the self? Does it imply rights for robots? Can machines be bad, dangerous, malicious, evil? [You will likely return to some of these questions in the fourth discussion, when you read Robot Ethics]
Second Discussion: Robotics Revolution, Science non-fiction

Leaving the fictional world, but not the speculative world, this set of texts aims to familiarize us with the world of robotics in the late 20th early 21st century. Rodney Brooks was the Panasonic Robotics Researcher at MIT from 1984-2008 and he served as head of the MIT Artificial Intelligence lab from 1997-2007 years. He is currently CEO of a robotics company, Rethink Robotics. This book, written for a popular audience in 2003, provides a good mix of history of artificial intelligence/robotics, a sense of the evolution of the field, as well as Brooks’ own innovations and creations in the field of robotics. In particular, he and his lab at MIT are largely responsible for early development of an action or behavior approach to robotics rather than a computational approach.

Much of what Brooks is discussing is hard to imagine if you haven’t seen it. Luckily, the internet and especially YouTube have dramatically increased what we can see, for free. I have included a few links below to footage of some of the earlier robots mentioned in the earlier chapters. Additionally, I have paired the Brooks reading with a selection of TED Talks that let us hear what researchers and ethicists are thinking about robots and see these emerging creations. Appendix A provides a longer (but not exhaustive) list of robot related TED Talks if you are interested.


Viewing: TED Talks

Discussion Questions:
• Chapters 1 and 2 provide overview for Brooks’ book and a quick overview of technological innovations that lead to this. What, according to Brooks, makes us human?
• What do we learn about humans or mechanical creatures from this particular set of robots/experiments?
  o Tortoises (W. Grey Walter)— https://www.youtube.com/watch?v=LULR1mXkKo
  o Shakey (Nils Nilsson)— https://www.youtube.com/watch?v=RhrLHkVuer
  o The cart (SAIL, John McCarthy) — http://web.stanford.edu/~learnest/cart.htm
• Chapters 3 and 4 provide accounts of different approaches and successes in robots from the mid-twentieth century to the early 21st. What do we learn about robots here? From a development perspective, what is most difficult to create and what is easier? Any surprises in these chapters?
Why according to Brooks are we interested in robots with human form? Do you agree?

How does Brooks define intelligence for this field? In what ways is this similar or different from how we typically think about human intelligence?

What is Kismet like? What can it do? What do we learn from social robots?

What difference does embodiment make according to Brooks?

Chapter 7 and 8 are speculative on Brooks’ part and his account of religion is not always generous, but here we get at some of the tension between what makes us special or not. Why does this notion seem to matter to us? How do robot discussions/imaginations trigger these fears. What does your group think about these issues? Consider rereading Genesis 1 and 2 in light of this conversation about robots. Also, remember the theology in Asimov’s story “Reason,” how does this inform this conversation?

The science fiction we read last session was clearly speculative and Asimov didn’t always accurately guess what would happen when. But this book by Brooks is now over ten years old. Have his predictions come true for the most part according to the TED Talks and your own knowledge of technology today? Where does he seem prescient and where does he seem off base?

What do the TED Talks add to the picture Brooks gives us?

Anything surprising or disturbing in the TED Talks?
Third Discussion: Empathy and fiction

For this third discussion we return to science fiction reading Philip K. Dick’s “Do Androids Dream of Electric Sheep” (1968) and watch Ridley Scott’s Blade Runner (1982/1994). This classic science fiction story is set in dystopian L.A. in our near future. Building on our conversation about social robots, Dick explores themes of empathy and life for humans, animals, and robots. I imagine many of you remember Ridley Scott’s 1982 film Blade Runner it was a blockbuster and became a cult classic. In 1994 Scott released a director’s cut with a significantly altered ending that is now considered the “real” Blade Runner. Be sure to read the novel before you watch/rewatch the film! I have also included a short reading about humans in relation to animals and robots for an optional framework for discussion.

Viewing: Blade Runner (1994)

Discussion Questions:

- Context: What is the earth like in this novel?
- What are the major issues raised by the androids? What are they like? How do humans interact with them? Along with androids, animals play an important role in this book. What does Dick seem to be implying about the value of life? Do the “specials” or “chicken heads” add to this picture of life?
- Last session you learned a bit about robots and emotions and how humans interact with them today. What role(s) does empathy play in this novel? Who has it, who doesn’t, what triggers it?
- What role does Mercerism play in the book? How does television via Buster friendly add to the story?
- In the book, is Deckard an “andy”? How does it change the story if he is or isn’t?
- The film pays more attention to science and the creation/life of replicants adding in a genetic cloning process. In particular the relationship between Roy Baty and Dr. Tyrell are emphasized in the film. Are there things that work better in the film than the book? Vice Versa? What about the relationship between Rick and Rachael in the film and book?
- In the film, is Deckard a replicant? How does it change the story if he is or isn’t?
- The key scene between Deckard and Rachael in the film is sometimes referred to in fan circles as “the rape scene”. What role does this scene play in the film? In the book?
Fourth Discussion: Ethics

In the past three sessions we have encountered a range of ethical issues addressed through fiction and development. This week we will sample current scholarship on robot ethics from across the fields of AI, computer engineering, law, philosophy, and sociology. For the most part these scholars are all thinking about the possibility of moral robots as well as what our responsibility to them might be. Thinking back to Asimov’s Laws of Robotics, and the comparison with tools, we don’t owe our tools anything, but we want them to work for us. If AI develops enough that robots (or computers for that matter) are conscious—can think, feel, remember—the game changes a bit as our films demonstrate poignantly. I have paired this set of readings with the film *Bicentennial Man*, which gets at the issue of rights and what we might owe our conscious creations.

The articles in this book were written to outline the major issues in a wide range of disciplines. Because these essays represent the “state of the field” they are excerpts and are not all easily accessible for a lay audience to follow. But you will get the main ideas from most articles. Feel free to read more or less than what I have assigned here. The essays are grouped thematically. I suggest you read Chapters 1-5 and 16-21. After that each group can choose where to focus their attention: read 7-9 on the military, 11-12 on the law, or 13-15 on psychology and sex. I have included a few discussion questions for most of the chapters below. But the goal is to learn about and respond to this emerging field.


**Viewing:** *Bicentennial Man* (1999)

In order to help lay some basic groundwork for this set of readings I have provided some very quick summaries of a few important concepts.

An overview of the three major ethical theories (quick and dirty version):

- **Deontological ethics**: Deontological ethics refers to rule or duty based ethics. Deontological ethics hold that the right thing to do is whatever adheres to the rules of ethics. In this theory of ethics the action itself is right if it follows the rules and wrong if it does not. The consequences can’t be used to judge the rightness of an action and the overall character of the doer is also inconsequential. Kant’s categorical imperative are typical examples of deontological ethics. [For a complete overview see: http://plato.stanford.edu/entries/ethics-deontological/ .]

- **Virtue ethics**: Most commonly associated with Plato and Aristotle, virtue ethics holds that right actions are performed by those with good character. Rather than relying on a clear set of rules or the consequences this theory of ethics places primacy on the character of the doer. We much cultivate ourselves to be virtuous and then our actions will be virtuous. [For a complete overview see: http://plato.stanford.edu/entries/ethics-virtue/ .]

- **Utilitarianism**: Probably the most widely known contemporary America, this theory is associated with Jeremy Bentham and John Stuart Mill. Sometimes referred to as consequentialism, utilitarian holds that the consequences are all we can use to judge the rightness of an action. Whatever maximizes the greatest good for the greatest number is
“good” or the right thing to do. [For a complete overview see: http://plato.stanford.edu/entries/consequentialism/]

The following two “tests” are designed to think through how we will know if a computer (AI) is thinking or conscious.

- **Turing test:** Alan Turing developed this test in an essay in 1950 as a way to determine when a computer can think. Turing posits the imitation game wherein two subjects and an interrogator are separated. The interrogator asks questions of each of the subjects to determine whether they are male or female. If a computer can fool the interrogator in such a test it would be able to think. Most contemporary discussions of the Turing test no longer rely on determining gender but whether a computer can “pass” for human in such a test. [Turing’s paper can be found here: http://loebner.net/Prizef/TuringArticle.html .]

- **Chinese Room test:** Philosopher John Searle’s Chinese room thought experiment is aimed at debunking the notion of the possibility of AI. In particular, it aims to prove that even if a computer could pass the Turing test it would not have consciousness. In this thought experiment there is a Chinese room. A person can send written Chinese into the room and get an answer back in Chinese. There could be a person who speaks/reads/understands Chinese in the room, or there could be a computer which just knows the language rules. Just because the computer spits back proper Chinese doesn’t mean it thinks.

Discussion questions:

- Chapters 1-4 lay out the basic issues that scholars and creators are thinking about today. What are the current trends in thinking about robots and ethics? What do we have to take into account?
- What role does fiction play in these conversations about robots today?
- Are there any issues raised in these introductory chapters that are startling? Rub you the wrong way? Are exciting?
- Chapter 3 by Keith Abney outlines modern theories of ethics, rights and responsibilities. Does approaching these issues through the lens of robots shed light on our notion of human ethics?
- Chapters 4 and 5 start add to our conversation about ethics through the perspective of design. What are the main issues that are raised in terms of responsibility? Allen and Wallach (chapter 4) are particularly well known for their concept of Moral Machines. How are they defining morality here? Is this a useful definition? Is this notion of moral machines persuasive to you or not? They even outline the major objections to their theory for you, are these persuasive?
- Japan has conducted significant robot research and many of today’s robots in practice are located there. Thus, there is a growing body of literature on the difference between Buddhist and Christian approaches to robot ethics. How does Hughes chapter on Buddhism add to your conversation about robots?
- Chapters 7-9 tackle the issue of robots in the military addressing major military ethics such as just war theory. Where are we today? What benefits can robots provide in
military situations? What are the risks? Does this conversation make you think differently about warfare?
  o See the TED Talk list in Appendix A for talks that address military robots.
- Chapters 11-12 address issues of robots in the law. This is also addressed in the film *Bicentennial Man* (and even more directly in the Asimov short story it is based on). What are the major issues that robots raise for the law.
- Chapters 13-15 address the issues of sex and psychology and robots. Thinking back to the fiction we read and earlier conversations about empathy and humans relating to robots, what are the ethical issues that arise here?
- Chapters 16-18 address the growing field of medical robots. What are the positives and negatives of this area of development? Robots are frequently seen as providing a kind of care for humans, what are the implications of this?
  o See the TED Talk list in Appendix A for talks that address military robots.
- Finally, chapters 19-21 address the issues of rights and responsibilities. In particular, sparrow’s Turing triage test adds a new twist on how we assess whether robots should be treated as persons with rights even if they aren’t human beings? Think through this test. How does it work? Is this distinction between a person and a human helpful? Persuasive? Beavers argues that our very concept of ethics might necessarily be changing if we pursue applying them to robots. What do you think? Here he addresses major ethical theories and the ways that they can or cannot be implemented in robots?
- What in the end does your group make of the notion of robots ethics, rights, our ethics towards robots? Think back to your discussion of Asimov’s three laws of robotics. Has the conversation shifted?
Fifth Conversation: Choices

I have left this final unit open ended, suggesting several clusters of films and/or readings that different groups can choose from. You are now familiar with the main terms, some fiction, and recent science and philosophy and can chose what topics your group would like to delve into further. Be sure to use this conversation as a way to wrap up the year’s themes.

Option 1: Danger! Losing ourselves
For groups who are up for more reading Sherry Turkle, also at MIT, started her career in the late 1980s and early 1990s positing the potential benefits of online life for participants. But in recent years, in part through encounters with Brooks and Brazeal’s robot labs, Turkle has come to argue that we are seriously losing track of our humanity in the wake of new media and new technology.


Option 2: Recent films
As I mentioned in the introduction to this syllabus, once you start watching for robots you will find they are everywhere. This year there were two significant major films featuring robots. Watch these recent films and compare them to mid/late 20th century fiction you read and the current robotics conversations.

**Viewing**: *Chappie* (2015) and *Ex Machina* (2015) [note: both are fairly violent]

Option 3: Children’s films
Think through the themes of this syllabus via these two animated films.

**Viewing**: *Wall-e* (2008) and *Big Hero 6* (2014)

Option 4: TED Talks Robots in the military

Watch the following TED Talk playlist: Will Drones save or destroy us
https://www.ted.com/playlists/135/will_drones_save_us_or_destroy

Option 5: TED Talks Robots and medicine

Watch the following TED Talks
https://www.ted.com/talks/henry_evans_and_chad_jenkins_meet_the_robots_for_humanity
https://www.ted.com/talks/catherine_mohr_surgery_s_past_present_and_robotic_future
https://www.ted.com/talks/eythor_bender_demos_human_exoskeletons
What Robots Teach Us About Being Human
CC Alumni Reading Group Syllabus 2015-16

Appendix A: TED Talks

How to live with Robots: TED Talk playlist
https://www.ted.com/playlists/17/how_to_live_with_robots

Will Drones save or destroy us: TED Talk playlist
https://www.ted.com/playlists/135/will_drones_save_us_or_destroy

Self-aware robots
https://www.ted.com/talks/hod_lipson_builds_self_aware_robots

Learning from robots
https://www.ted.com/talks/ken_goldberg_4_lessons_from_robots_about_being_human

Rodney Brooks
https://www.ted.com/talks/rodney_brooks_why_we_will_rely_on_robots
https://www.ted.com/talks/rodney_brooks_on_robots

War
https://www.ted.com/talks/pw_singer_on_robots_of_war
https://www.ted.com/talks/daniel_suarez_the_kill_decision_shouldn_t_belong_to_a_robot
https://www.ted.com/talks/eythor_bender_demos_human_exoskeletons

Medicine
https://www.ted.com/talks/henry_evans_and_chad_jenkins_meet_the_robots_for_humanity
https://www.ted.com/talks/catherine_mohr_surgery_s_past_present_and_robotic_future
https://www.ted.com/talks/eythor_bender_demos_human_exoskeletons