Ethical Foundations

CS 4002 – Robots and Society
So who’s right?

- Bill Joy?
- Bill Gates?
- Stephen Hawking?
- All?
- Neither?

Joy

- “Biological Species never survive encounters with superior competitors”
- “Biological humans would be squeezed out of existence”
- “We are on the cusp of the further perfection of extreme evil”
- “We would be more secure if we did not pursue them” [GNR] (relinquishment)
Gates

• “A Robot in Every Home”
• “Robotic devices will become a nearly ubiquitous part of our day-to-day lives.”
• “these devices will have a profound impact on the way we work, communicate, learn, and entertain ourselves as the PC has had over the past 30 years”.

Hawking

• “Tempting to dismiss the notion of highly intelligent machines … would be… potentially the worst mistake in history”
• “Success in creating AI would be the biggest event in human history. Unfortunately it may also be the last.”
• “No fundamental limits to what can be achieved”
For Sale (Joy’s/Hawking’s vision realized):

Robot Vacuum Eats Sleeping Woman’s Hair

A woman whose hair was swallowed up by a robot vacuum cleaner while she was sleeping on the floor had to be freed by emergency crews.

The 52-year-old, called Yoon, woke in agony at her South Korean home after the robot apparently mistook her hair for fluff.

Unable to free herself, she called the Changwon city fire department in desperation.

Four emergency workers took half an hour to free her after opening the device.

The woman escaped serious injury.
Why Roboethics?

• Let’s start with null hypothesis
  
  – Is it different than any other form of ethical studies?
  
  – Does it warrant a course separate from “Computers and Society?”

Why such a forum?

• Ethics are essential for professional conduct
• Robots are penetrating society at all levels
  – Home/Service
  – Entertainment
  – Military
  – Prostheses
  – What else? space, manufacturing
• These discussions will enable you to understand ethical consequences of your work, and to (hopefully) act in a socially-responsible manner
Norbert Wiener as an ethicist

Father of Cybernetics (1894-1964)

"Those of us who have contributed to the new science of cybernetics...stand in a moral position which is...not very comfortable. We have contributed to the initiation of a new science which...embraces technical developments with great possibilities for good and for evil. We can only hand it over into the world that exists about us, and this is the world of Belsen and Hiroshima"
More Wiener quotes

“By the very slowness of our human actions, our effective control of our machines may be nullified.”

“If we want to live with the machine…we must not worship the machine. We must make a great many changes in the way we live with other people….We must turn the great leaders of business, of industry, of politics, into a state of mind in which they will consider…people as their business and not as something to be passed off as none of their business.”

“The lords of the present science…are nothing more than apprentice sorcerers, fascinated with the incantation which starts a devilment that they are totally unable to stop. Even the new psychology…becomes in their hands a way for obliterating the conscientious scruples of the working scientists, and for destroying such inhibitions as they may have against rowing into this maelstrom….

“Let these wise men who have summoned a demoniac sanction for their own private purposes remember that in the natural course of events, a conscience which has been bought once will be bought twice. The loyalty to humanity which can be subverted by a skillful distribution of administrative sugar plums will be followed by a loyalty to official superiors lasting just so long as we have the bigger sugar plums to distribute.”
One more Wiener quote:

“The relations of the machine to the living being, and with systems involving elements of both kinds [present] some of the most important moral traps into which the present generation of human beings is likely to fall.”

—Norbert Wiener, *God & Golem, Inc*

Fuel for the fire follows:
Cockroach Controlled Mobile Robot

Baby Robot
Meet the xenobot: world's first living, self-healing robots created from frog stem cells

Kevin Warwick and His wife, Irena
Some applications:

(From Discovery Channel)
Entertainment: AIBO

Paro
Dance Partner Robot
Kosuge Lab - Japan

Controversies

- Robots as our Mind Children (Moravec/Kurzweil)
- Robot Slave Society (Albus/Bryson)
- Robots as partners (Ishiguro/Arkin)
- Robot-human hybrids (Clark/Warwick)
- Fatalism/Extinction/Relinquishment – Joy
- Neo-Luddism - Unabomber
What is ethics?

- Philosophical study of morality, where morality is the set of rules determined by society
- Ethics is based in reason, hence explaining why things are the way they are or should be is important
- Form the standards of conduct and moral judgment
- The fact that right and wrong exist serves as the underlying assumption.
- Can operate at both societal and personal levels
- Basically, answering “what is the right thing to do? Why?”
- Valid logical argumentation is central to convince others

Why study ethics?

- ?
- Should there be ethical theories in the first place?
- Can universal ethics be achieved?
- Are morals simply a matter of opinion? Or are they based on some absolute?
- What is duty/responsibility in presence of unethical acts?
Is Ethics Scientific?

- ?
- It is based on rationality
  - Based on rational moral principles
  - Uses sound, carefully reasoned arguments
- It is not as objective as science
- Although no unique correct solution guaranteed, alternatives can often be weighed

Possible Costs/Benefits of an Ethical System

- ?
- Costs
  - Prohibition of certain actions by anyone
  - Obligatory actions for everyone
  - Enforcement/punishment system required
- Benefits
  - Shared responsibilities
  - Increase in social productivity / quality of life
  - Security against external dangers
Early Ethics

• Socrates (2400 years ago) described ethical reasoning, when Plato wrote in Crito Dialogue the ethics of Socrates’ imprisonment and death as opposed to flight into exile

Virtue Ethics

• Plato and Aristotle
• Resurgence in 1950s  
  (Anscombe’s Modern Moral philosophy)
• Based on individual character rather than rules  
  – Virtue: character trait to do that which is good  
  – Practical wisdom (phronesis): the knowledge to know what to do that’s right  
  – Eudaimonia: flourishing, happiness, or well-being.  
  Virtue is considered necessary for this state
Objections to Virtue Ethics

- Action-guidance problem
  - Non-codifiable: cannot represent logically
  - Non-virtuous person cannot apply this form of ethics or even understand it
- Conflict Problem
  - How to deal with ethical dilemmas?
- “Concerned with being rather than doing”
- Egoistic – centered on the self

Ethical Theories

- Act in the manner that is for the best (Consequentialism)
  - Examples:
    - If for oneself: Egoism
    - If for everyone: Utilitarianism
- Conform to the moral law (Rights-based or Deontology)
  - Examples:
    - Derived from culture (Relativism)
    - From God’s law (Divine command)
    - Universal logic and reason (Kant)
Subjective Relativism

- No absolute right or wrong, not even a single standard
- Culture and time determine your moral standards
- If it feels right to you, then it is ok.
- Everyone decides what is right or wrong for themselves. “What’s right for you may not be right for me”

Egoism

- Related to subjective relativism
- Pure self interest
- Some forms of Libertarianism (E.g, Ayn Rand)
- Not only at personal level, but also
  - National (U.S. policy?)
  - Corporate (Enron?)
  - Species (Robo Sapiens?)
Cultural Relativism

- Still, there is no absolute right or wrong
- Culture and time determine moral standards
- Those in power decree moral norms
- Societies are basis for what is acceptable or not (social conventions) – right and wrong is based on societal moral guidelines.
- Morality varies over time and place to place

Justifications for Cultural Relativism

- It is inappropriate or arrogant for one society to impose its own morality on or judge another society
- Different contexts require different moral guidelines – the world changes, so must morality (e.g., advent of robots)
- Morality is rooted in behavior – looking at how people behave is a basis for morality
Arguments against cultural relativism

- Differing moral systems can be evaluated, the best ones should be imposed on other societies
- How are the guidelines determined? What is their basis?
- Lack of absolute provides no basis for reconciling conflicting cultures
- Some core values are shared across societies irrespective of time and place (e.g., telling truth to each other, babies must be taken care of, prohibition against murder, etc.)

Divine Command Theory

- Religious Traditions – (e.g., Jews, Christians, and Muslims, all acknowledge the Torah (Books of Moses))
- Good actions of individual are consistent with will of God, bad actions are against will.
Divine Command Theory

- Justifications for
  - God is all knowing and all-good
  - Obedience to our Creator is important/required
  - Ultimate authority provides morality

- Arguments against
  - Many different religions/holy scriptures (even versions of same religion’s book)
  - Secular society will likely resist adoption of a strict religion-driven morality
  - Not all morality addressed in religious literature (robots?)
  - Divine theory is based on obedience not reason

In-class Exercise
Utilitarianism

- Based on the principles of moral consequences
- Person should act to increase overall benefit/utility for all
- The good of the many outweighs the good of the few
- An individual may suffer severe consequences as the basis for his/her actions

Act Utilitarianism

- Bentham and Mill (18-19th century English philosophers)
- Principle of Utility:
  - An action is right or wrong to the extent that it increases or decreases the total happiness of the affected parties.
- Utility is the ability of an object to produce happiness/unhappiness in a community or individual
- Motive is unimportant – consequences are
- Complete analyses are required which may be complex – it is even an open question as to who to include in the “affected parties” (robots?)
Act Utilitarianism

- Justifications for
  - Happiness is central: if a purpose of life is to be happy than the Principle of Utility addresses that
  - Straightforward and grounded in the here and now
  - Comprehensive – considers broad situation
- Arguments against
  - Although a calculus is prescribed it is not clear exactly how to apply it and to whom
  - Hard to do since it is complex and time consuming. Does not fit well with everyday decision making.
  - Ignores a person’s sense of duty (innate?) and seems counter-intuitive
  - Concept of moral luck involved (both good and bad) – e.g., unintended adverse results of otherwise good willed actions

Rule Utilitarianism

- Followed on the heels of act utilitarianism to address its potential flaws
- Instead of action-based, it is rule based.
- Adopt moral rules that are consistent with the Principle of Utility, i.e., they maximize happiness.
- Rules should be followed without exception
  - same as Kant, but there the categorical imperative (moral principle that behavior should be determined by duty) is driver, not the Principle of utility
Rule Utilitarianism

• Justifications for
  – Easier to evaluate than act utilitarianism
  – Rules can be relied on for everyday acts
  – Moral rules can be steadfast in light of an exception (promise breaking ok, contrary to Kant)
  – Moral luck not an issue, as rule abstracts away from action

• Arguments against all utilitarianism
  – Single scale to evaluate all kinds of consequences (which may not be easily comparable)
  – Unjust distribution of good consequences is a possible problem

One Ethical Analysis Framework

1. What is your moral intuition regarding the issues of the case (regarding right or wrong)?
2. Develop a coherent rationale that defends your position
   • Does it optimize greatest utility for all parties?
   • Does it violate any ethical duties?
   • If conflict exists, what is the higher duty?
3. What are the implications on other stakeholders and public policy in general?
Stakeholders

- ?
- Who’s involved?
  - Actor
  - Society
  - Nation
  - Company
  - Environment
  - Robots
  - Religious/governmental institutions
  - Future generations
  - Others?
- Stakeholder analysis – consider other viewpoints

Case Study I

Use of Robots as babysitters for children
(Robot Nanny)
- What are your intuitions?
- Who are stakeholders?
- What are issues and implications?
- What if any moral guidelines should be formulated?
- What about robots as caretakers for elderly?
Deontology

- Every person warrants respect and has worth
- Motivation and intent are basis for morality
- Moral law, promises, and commands are sources of obligations
- E.g., Declaration of Independence “All men are created equal. With certain inalienable rights”.
- Ethics is based on appropriate decisions on who and how to curtail rights of those outside of moral norms.

Universal logic and reason (Kant)

- Immanuel Kant 18th Century German philosopher
- Duty-based ethical system
- People should be guided by moral laws that are universal
- Morality must be based on reason
- Instead of simply quoting scripture as basis for moral law, reason can provide an explanation
- Good will is the basis for positive moral actions and is the only universally good thing.
- What we want to do is of no importance, what we ought to do is. This is referred to as dutifulness.
- A moral value of an action is based on its relationship to an underlying moral rule.
Categorical Imperative (Kant)

- “Act only from moral rules that you can at the same time will to be universal moral laws” (e.g., is promise-breaking acceptable? If everyone breaks promise, then there is no such thing as a promise, hence it is unacceptable to do so.)
- “Act so that you always treat both yourself and other people as ends in themselves, not as a means to an end.”

Universal logic and reason (Kant)

- Justifications for
  - Rational
  - Generates universal moral guidelines
  - All humans are treated morally equally
- Arguments against
  - Single rules may not completely cover a situation (stealing food to feed your children?)
  - No method to resolve conflict between rules.
  - Real world resists axiomatization
  - No exceptions to moral laws are allowed
Social Contract Theory

“Morality consists in the set of rules, governing how people are to treat one another, that rational people will agree to accept, for their mutual benefit, on the condition that others will follow the rules as well”

[Rachels 03]

17th century English Philosopher Thomas Hobbes, later 18th Century Swiss philosopher Rousseau

All rights belong to the whole community – community determines rules and all members must adhere.

No one is above the law, so rules must be fair

People agree to follow the moral rules under the belief that all will follow them.

People have rights:
- Negative rights: they can choose to exercise (speech)
- Positive rights: that society must provide (free education)

Principles of justice prescribe notion of free and equal citizens
- allows for redistribution of wealth through differential taxation and services
- Allows for opportunity for social or economic advancement within society
Social Contract Theory

• Justifications for
  – Uses rights as the basis for morality. This is often attractive to many cultures
  – Explains why self-interest occurs by rational people in absence of social contract
  – Relates ethical issues between people and government

• Arguments against
  – No one agreed to the contract – it’s not a real entity.
  – Actions can be interpreted in multiple ways
  – Does not solve moral problems when rights are in conflict (e.g., pro-life, pro-choice)
  – May be unjust to those incapable of holding up to contract (drug addicts, mentally ill, children)

Rawls’ Theory of Justice

• Main Principles
  – Each person may claim a fully adequate # of rights and liberties as long as these are consistent with everyone else having the same claim.
  – Social and economic inequalities are acceptable if:
    • They are associated with positions anyone with equal capability can assume (independent of socio-economics)
    • They are to the greatest benefit of the least advantaged (redistribution of wealth towards equality for all)
Duty-Based Ethics (Ross)

Basic moral duties that are indisputable and self-evident

1. Fidelity: Keep promises and tell the truth
2. Reparation: Right wrongs inflicted on others
3. Justice: Distribute goods equitably
4. Beneficence: Improve the lot of others for virtue, intelligence, and happiness
5. Self-improvement: Improve your own intelligence and virtue
6. Gratitude: Be thankful where appropriate
7. Non-injury: Avoid hurting others

Feminist Ethics

Justifications For (Pros)

• Arose from concern that traditional Western ethics devalues women
  – Ignore or trivialize virtues of character commonly associated with women
• “Communal woman” rather than “autonomous man”
• Maternal approaches
  – Fostering, training, attentive love
• Political:
  – concerned with power, rather than absolutes of good and evil.
  – How to eliminate subordination of women.
  – Focus on women-oppressive system and structures
  – Achieve gender equity
Objections to Feminist Ethics

- While origins of traditional ethics may have been male-centered, they can be gender neutral
- Doubt that any one relationship should serve as paradigm for all human relationships
- Differences in mothering, including over-idealizing, potential for abuse and neglect prevent generalization

Where do these instances fit?

- The Golden Rule
- What is fair for one, is fair for all
- All’s fair in love and war
- Rights of minorities
- Royal decrees
- Dr. Phil
What is an Ethical Life?

• Excessive self-interest is often basis for unethical behavior
• Define/adopt high standards of ethical conduct and live by them
• Stand by your standards even at personal cost
• Adopt vision of a larger purpose in life
• Accept personal responsibility for decisions
• Consider ethical consequences when they appear
• Avoid situations that compromise ethics

Applied Ethics – Living in the Real World

• No single theory universally followed
• Broad disagreement on what’s right and wrong – individualism, religion, culture
• Tolerance of other points of view
• Ideal action is not always clear-cut in differing situational contexts. (examples?)
• Goal may be to follow at least a minimal ethical standard – where is the line drawn for acceptable behavior of yourself and others
Professional Ethics – Living within your career

- Codes of conduct (ACM, IEEE, etc.)
- Relate to relationships between professionals and their clients, society, other professionals, employees, employer
- Specifics of a particular profession (examples?)

Some ethical questions from a robotics perspective

- Can a robot have a religion?
- Can a person marry a robot?
- Can a person enslave a robot?
- Can a robot be abused?
- Can a person turn off/reset a robot?
- Asimov’s laws: right or wrong?
One Ethical Analysis Framework - Revisited

1. What is your moral intuition regarding the issues of the case (regarding right or wrong)?

2. Develop a coherent rationale that defends your position
   • Does it optimize greatest utility for all parties?
   • Does it violate any ethical duties?
   • If conflict exists, what is the higher duty?

3. What are the implications on other stakeholders and public policy in general?

Stakeholder Analysis

1. Identify Stakeholders
2. Identify utility for each
3. Assign priorities
4. Assess impact of action on all parties
5. Weigh results
   and
   What rights of any stakeholders are violated?
Case Study II

• Your close relative (child, sibling) announces they want to marry a robot.
  – What are your intuitions?
  – Who are stakeholders?
  – What are issues and implications?
  – What if any moral guidelines should be formulated?

Case Study III

• Can a robot be turned off arbitrarily?
  – Every day lab robot
  – Hal 9000, Bicentennial Man, R2D2, etc.
  – Is there a transition point?
  – What are your intuitions?
  – Who are stakeholders?
  – What are issues and implications?
  – What if any moral guidelines should be formulated?